AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXVIII, Number 3
JULY, 1956

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

The American Journal of Mathematics appears four times yearly.

The subscription price of the Journal is $8.50 in the U. S.; $8.75 in Canada; and $9.00 in other foreign countries. The price of single numbers is $2.50.

Manuscipts intended for publication in the Journal should be sent to Professor Aurel Wintner, The Johns Hopkins University, Baltimore 18, Md.

Subscriptions to the Journal and all business communications should be sent to The Johns Hopkins Press, Baltimore 18, Maryland, U. S. A.

The Johns Hopkins Press supplies to the authors 100 free reprints of every article appearing in the American Journal of Mathematics. On the other hand, neither The Johns Hopkins Press nor the American Journal of Mathematics can accept orders for additional reprints. Authors interested in securing more than 100 reprints are advised to make arrangements directly with the printers, J. H. Furst Co., 20 Hopkins Place, Baltimore 1, Maryland.

The typescripts submitted can be in English, French, German or Italian and should be prepared in accordance with the instructions listed on the inside back cover of this issue.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

ARTIN-SCHREIER EQUATIONS IN CHARACTERISTIC ZERO.*

By R. E. MacKenzie and G. Whaples.

The proof of the existence theorem of generalized local class field theory [6, 7] shows that if k is a field over which that theory holds and p is the characteristic of its residue class field then the cyclic extensions of k of degree p fall into two disjoint classes. Each extension of the first class is contained in the composite of finitely many extensions generated by root of Artin-Schreier equations

$$x^p - x - \lambda = 0, \qquad \lambda \varepsilon k, \qquad | p^p \lambda^{p-1} | < 1. \tag{1}$$

Each extension of the second class is generated by a root of an equation

$$x^p - \alpha = 0, \qquad \alpha \varepsilon k, \qquad | \alpha | \not\equiv | k^{\cdot p} |, \tag{2}$$

where $| k^{\cdot p} | = \{ | \beta^p | \; | \; \beta \varepsilon k^{\cdot} \}$. Extensions of the second class occur only when k has characteristic zero and contains primitive p-th roots of unity.

This suggests the conjecture that every extension of the first class is generated by a root of *one* Artin-Schreier equation. This is well known, of course, when k has characteristic p [9]. In this paper we prove the conjecture not only for the fields of local class field theory but for all fields which are maximally complete [2, 4] under an arbitrary non-archimedean valuation (not necessarily discrete or of rank one) and which have a residue class field with no inseparable extension. Even without this assumption on the residue class field our methods give considerable information about the cyclic extensions of degree p.

In local class field theory the two classes of extensions can be distinguished by certain inequalities involving either the conductor or the different. In our more general situation the conductor and different are unavailable but we introduce another invariant, the distortion constant, which takes their place and happens also to simplify the computations very much.

This gives a method of assigning to each cyclic extension of degree p a canonical defining equation. In the local class field case this poses the

* Received May 26, 1954; revised February 3, 1956. Work of Whaples on this paper supported in part by National Science Foundation Grant NSF G1916.

problem of finding an explicit reciprocity law, namely, a rule translating the class field theory parametrization of extension fields and their automorphisms into the parametriaztion given by our canonical defining equations. Such laws have up to now been restricted to the case where k has characteristic p or contains primitive p-th roots of unity due to the lack of a canonical defining equation in the other case. The problem of finding an explicit reciprocity law separates naturally into two parts: (A) Find an explicit formula for the norm residue symbol defined on some particular basis for elements of $k^\cdot$ modulo norms. (B) Find an explicit basis for the group of norms. We have solved (A) but not (B).

Oystein Ore [3] has also studied generating equations for such cyclic extensions (not only of degree p but of degree p^n) in the absence of p-th roots of unity. Although his general ideas are similar to ours he restricts himself to algebraic number fields, uses congruences instead of equations, and, because he wants congruences with integral coefficients, has several standard forms instead of only two.

For brevity we state at the beginning of each section the assumptions made in it and do not repeat these assumptions in stating propositions and theorems.

1. Orthobases. Let K have a nonarchimedean valuation $|\ |$ with no assumptions of completeness. Let K/k be finite algebraic and $\bar{K}/\bar{k}$ the residue class field extension. Let n, e, f denote the degree of K/k, the ramification number of K/k, and the degree of $\bar{K}/\bar{k}$, respectively.

Definition 1. A sequence $\mathrm{A}_1, \mathrm{A}_2, \cdots, \mathrm{A}_n$ of elements of K is called an *orthobasis for* K/k (relative to $|\ |$) if it is a basis and

$$|\textstyle\sum_\nu \alpha_\nu \mathrm{A}_\nu| = \max_\nu |\alpha_\nu \mathrm{A}_\nu| \text{ for all } \alpha_1, \alpha_2, \cdots, \alpha_n \in k. \tag{3}$$

The name was chosen because, like an orthonormal basis of a normal vector space, an orthobasis makes absolute values easy to compute. If the valuation is not discrete there may not exist any minimal basis for K-integers over k-integers, but an orthobasis is an excellent substitute.

PROPOSITION 1. *K/k has an orthobasis if and only if $ef = n$.*

Proof. Let $\{\mathrm{A}_i\}$ be a finite set of elements of K such that $\{|\mathrm{A}_i|\}$ are in distinct cosets modulo $|k^\cdot|$ and let $\{\mathrm{B}_j\}$ be elements of value 1 whose residue classes are linearly independent over $\bar{k}$. It is easy to see that the elements $\{\mathrm{A}_i\mathrm{B}_j\}$ satisfy (3) and are therefore linearly independent over k. The well known statement $n \geqq ef$ [1] follows. If $n > ef$ then any set of n

elements of K contains more than f elements whose values are in the same coset modulo $|k^\cdot|$ and hence fails to satisfy (3).

Definition 2. An element $\mathrm{A} \varepsilon K$ is called an *orthogenerator for* K/k if $1, \mathrm{A}, \mathrm{A}^2, \cdots, \mathrm{A}^{n-1}$ are orthobasis.

2. Distortion constant. Besides the assumptions of Section 1 let K/k be cyclic with a generating automorphism σ such that

$$|\sigma \mathrm{A}| = |\mathrm{A}| \text{ for all } \mathrm{A} \varepsilon K. \tag{4}$$

Definition 3. An element $\Gamma \varepsilon K$ is called a *distortion constant* (d.c.) for K/k when

$$\sigma \mathrm{A} = \mathrm{A}(1+\Gamma) \text{ for some } \mathrm{A} \varepsilon \mathrm{K} \text{ and} \tag{5}$$

$$\sigma \mathrm{B} = \mathrm{B}(1+O(\Gamma)) \text{ for all } \mathrm{B} \varepsilon K, \tag{6}$$

where $O(\Gamma)$ denotes an element of value $\leqq |\Gamma|$.

Clearly K/k has a d.c. if and only if $((\sigma - 1)\mathrm{B}/\mathrm{B} = \mathrm{B}^{\sigma-1} - 1$ assumes a maximum value for some $\mathrm{B} = \mathrm{A} \varepsilon K$. If any d.c. exists the value of every d.c. equals this maximum, hence is an invariant of K/k and σ. Since $\Gamma = ((\sigma - 1)\mathrm{A})/\mathrm{A} = \mathrm{A}^{\sigma-1} - 1$ it follows from (4) that $|\Gamma| \leqq 1$. By induction using (5) we obtain for all i

$$\sigma^i \mathrm{A} = \mathrm{A}(1+\Gamma)(1+\sigma\Gamma)\cdots(1+\sigma^{i-1}\Gamma) = \mathrm{A}(1+O(\Gamma)). \tag{7}$$

Similarly from (6) we see that $\sigma^i \mathrm{B} = \mathrm{B}(1+O(\Gamma))$ for all i and all B. If τ is another generating automorphism and Γ' a d.c. corresponding to it then $\Gamma' = O(\Gamma)$ and $\Gamma = O(\Gamma')$ so $|\Gamma|$ does not depend on the choice of σ and is an invariant of K/k.

The essential idea of a d.c. and of the following proposition appears in O. F. G. Schilling's book [4, pp. 80, 81].

PROPOSITION 2. *If K/k has an orthobasis $\mathrm{A}_1, \mathrm{A}_2, \cdots, \mathrm{A}_n$ and $\mathrm{A}_i^{\sigma-1} - 1$ is of maximal value for $i = 1, 2, \cdots, n$ then it is a d.c. If A is an orthogenerator then $\mathrm{A}^{\sigma-1} - 1$ is a d.c.*

Proof. Let $\mathrm{A}_1, \mathrm{A}_2, \cdots, \mathrm{A}_n$ be an orthobasis and let $\Gamma = \mathrm{A}_i^{\sigma-1} - 1$ be of maximal value. Then for all $\mathrm{B} = \sum \alpha_\nu \mathrm{A}_\nu \varepsilon K$ we have $\sigma \mathrm{B} = \sum \alpha_\nu \mathrm{A}_\nu (1 + O(\Gamma))$ so $|(\sigma - 1)\mathrm{B}| \leqq \max_\nu |\alpha_\nu \mathrm{A}_\nu O(\Gamma)| \leqq |\mathrm{B}\Gamma|$ by (3). Hence $\sigma \mathrm{B} = \mathrm{B}(1+O(\Gamma))$ and Γ is a d.c. If A is an orthogenerator and $\sigma \mathrm{A} = \mathrm{A}(1+\Gamma)$ then $\sigma \mathrm{A}^\nu = \mathrm{A}^\nu (1+\Gamma)^\nu = \mathrm{A}^\nu (1 + O(\Gamma))$ because $|\Gamma| \leqq 1$. So from what was just proved Γ is a d.c.

3. Cyclic extensions of degree p. Besides the assumptions of Sections 1 and 2 assume K/k is cyclic of degree p where $0 \leqq |p| < 1$ and $ef = p$. p is the characteristic of $\bar{k}$ and may or may not be that of k. Our results are of interest only in the case when k has characteristic 0.

Then K/k always has an orthogenerator $\mathbf{A}$.

Definition 4. If $e = p$ let $\mathbf{A}$ be any element with $|\mathbf{A}| \not\leqq |k\cdot|$. If $f = p$ and $\bar{K}/\bar{k}$ is inseparable let $\mathbf{A}$ be any element of value 1 whose residue class is not in $\bar{k}$. If $f = p$ and $\bar{K}/\bar{k}$ is separable choose $\mathbf{A}$ of value 1 so that $\sigma\bar{\mathrm{A}} = \bar{\mathrm{A}} + 1$, i.e., $\sigma\mathbf{A} = \mathbf{A} + 1 + o(1)$. This is possible because $\bar{k}$ has characteristic p [8]. Call an $\mathbf{A}$ chosen in this way a *canonical orthogenerator* of K/k.

K/k has a d.c. Γ, and $|\Gamma| < 1$ unless $\bar{K}/\bar{k}$ is separable of degree p. We propose now to obtain sharp estimates of the relative values of $\mathbf{B}$, $(\sigma - 1)\mathbf{B}$, and $S\mathbf{B}$ of the sort usually obtained by using the different.

Let $\mathbf{A}$ be a canonical orthogenerator and $\sigma\mathbf{A} = \mathbf{A}(1 + \Gamma)$. Then $(\sigma - 1)\mathbf{A}^i = \mathbf{A}^i((1 + \Gamma)^i - 1)$. If $|\Gamma| < 1$ then $(1 + \Gamma)^i - 1 = i\Gamma + o(\Gamma)$ and has value equal to $|\Gamma|$ when $p \nmid i$. If $|\Gamma| = 1$ then, according to definition 4, $\bar{\Gamma} = \bar{\mathrm{A}}^{-1}$ so $\bar{\Gamma}$ generates $\bar{K}/\bar{k}$ and $|(1 + \Gamma)^i - 1| = 1 = |\Gamma|$ when $0 < i < p$. This proves

$$|(\sigma - 1)\mathbf{A}^i| = |\Gamma\mathbf{A}^i| \text{ for } 1 \leqq i \leqq p - 1 \tag{8}$$

in all cases. Consider the elements

$$1, \Gamma^{-1}(\sigma - 1)\mathbf{A}, \Gamma^{-1}(\sigma - 1)\mathbf{A}^2, \cdots, \Gamma^{-1}(\sigma - 1)\mathbf{A}^{p-1}. \tag{9}$$

If $|\Gamma| < 1$ then the previous remarks show that these element are equal to $1, \mathbf{A} + o(\mathbf{A}), 2\mathbf{A}^2 + o(A^2), \cdots, (p - 1)\mathbf{A}^{p-1} + o(\mathbf{A}^{p-1})$, respectively, and hence form an orthobasis. If $|\Gamma| = 1$ then $\bar{K}/\bar{k}$ is separable of degree p and $\bar{\Gamma} = \bar{\mathrm{A}}^{-1}$. The residue classes of the elements (9) are $\bar{\mathrm{A}}^{i+1}((1 + \bar{\mathrm{A}}^{-1})^i - 1)$ for $i = 1, \cdots, (p - 1)$. They form a basis of $\bar{K}/\bar{k}$ so the elements (9) form an orthobasis of K/k. It therefore follows that the elements

$$\Gamma, (\sigma - 1)\mathbf{A}, (\sigma - 1)\mathbf{A}^2, \cdots, (\sigma - 1)\mathbf{A}^{p-1} \tag{10}$$

form an orthobasis in all cases.

Let $\mathbf{B}$ be any element of K. If we express $\mathbf{B}$ in terms of the orthobasis (10) we obtain $\mathbf{B} = \beta\Gamma + \Delta$ where $\beta \varepsilon k$, $S\Delta = 0$, and $|\beta\Gamma| \leqq |\mathbf{B}|$. Since $S\mathbf{B} = \beta S\Gamma$ we obtain

$$|S\mathbf{B}| \leqq |\Gamma^{-1}S\Gamma|\,|\mathbf{B}| \tag{11}$$

and equality holds only if $|\beta\Gamma| = |\mathbf{B}|$, that is, $|\mathbf{B}|\,\varepsilon\,|\Gamma|\,|k\cdot|$.

We may also express B in terms of the powers of a canonical orthogenerator A; $B = \alpha_0 + \alpha_1 A + \cdots + \alpha_{p-1}A^{p-1}$. Then $(\sigma - 1)B$ is expressed in terms of the orthobasis (10). Using (8) we obtain

$$|(\sigma-1)B| \leqq |\Gamma B| \tag{12}$$

and if we put $B' = B - \alpha_0$ we can say there is a B′ such that

$$(\sigma-1)B' = (\sigma-1)B \text{ and } |B'| = |\Gamma^{-1}(\sigma-1)B|. \tag{13}$$

THEOREM 1. *$|\Gamma^{p-1}| \geqq |p|$ for all K/k satisfying the assumptions of Section 3.*

Proof. Let A be a canonical orthogenerator and Γ its d.c. We can assume $|\Gamma| < 1$. By (7) we see that $\sigma^p A = AN(1+\Gamma)$ so $N(1+\Gamma) = 1$, i.e. $S\Gamma + E_2 + \cdots + E_{p-1} + N\Gamma = 0$ where E_i are the elementary symmetric functions. Now $|S\Gamma| < 1$. Since $|\Gamma| < 1$ it follows from (11) that $S\Gamma^i = o(S\Gamma)$ for $i > 1$ so, by Newton's identities, $E_2 + \cdots + E_{p-1} = o(S\Gamma)$ and

$$N\Gamma = -S\Gamma(1 + o(1)). \tag{14}$$

Substituting $B = 1$ in (11) we get $|p| \leqq |\Gamma^{-1}S\Gamma|$, i.e. $|S\Gamma| \geqq |\Gamma p|$. So (14) gives $|N\Gamma| = |\Gamma^p| \geqq |\Gamma p|$ and Theorem 1 follows.

If k has characteristic 0 then it is not possible without completeness assumptions to prove that K/k is generated by a root of an equation of either of the types (1) or (2) (see appendix). The best we can do is to derive, in Theorem 2 below, conditions that there exist an element satisfying a congruence related to (1).

Definition 5. Let $\wp(x)$ denote the polynomial $x^p - x$. If k has characteristic 0 let $\phi(x, y)$ be the polynomial with integal coefficients such that $(x+y)^p = x^p + y^p + p\phi(x, y)$.

Note that if $|A| \geqq |B|$ then $\phi(A, B) = O(A^{p-1}B)$ and that

$$\wp(x+y) = \wp(x) + \wp(y) + p\phi(x, y).$$

THEOREM 2. *Let K/k satisfy the assumption of Section 3. The following three conditions are equivalent: K contains an element Λ with*

$$(\sigma-1)\Lambda = 1 + o(1); \tag{15}$$

K contains an element M with

$$|pM| < 1 \text{ and } SM = -1; \tag{16}$$

$$|\Gamma^{p-1}| > |p|. \tag{17}$$

If Λ *satisfies* (15) *then* $|\Lambda| \geqq |\Gamma^{-1}|$ *and* $\Lambda = \alpha + \Lambda'$ *where* $\alpha \varepsilon k$, Λ' *satisfies* (15), *and* $|\Lambda'| = |\Gamma^{-1}|$. *For any such* Λ'

$$(18) \qquad (\sigma - 1)\wp(\Lambda') = o(1);\ \wp(\Lambda') = \lambda + o(\Lambda'), \lambda \varepsilon k, |\lambda^{p-1}| < |p^{-p}|.$$

If K/k *is ramified, conditions* (15), (16), (17) *hold if and only if* $|\Gamma| \not\leqq |k^{\cdot}|$.

Proof. The first three conditions are true when k has characteristic p so we may assume k has characteristic 0. If Λ satisfies (15) we can write $(\sigma - 1)\Lambda = 1 + p\mathrm{M}$ with $|p\mathrm{M}| < 1$. Then $0 = p + pS\mathrm{M}$ so M satisfies (16). If M satisfies (16) and $\Lambda = \sum_{1}^{p-1} \nu\sigma^{\nu}\mathrm{M}$ then Λ satisfies (15). So the first two conditions are equivalent. From (11) we see that $-\Gamma/S\Gamma$ is an element of minimal value with trace -1 so (16) is solvable if and only if $|p\Gamma| < |S\Gamma|$, which by (14) is equivalent to $|\Gamma^{p-1}| > |p|$. So the three conditions are equivalent.

Now let A be a canonical orthogenerator for K/k. Let Λ satisfy (15) and let $\Lambda = \alpha_0 + \alpha_1\mathrm{A} + \cdots + \alpha_{p-1}\mathrm{A}^{p-1}$. Then $\Lambda' = \Lambda - \alpha_0$ satisfies (13). Hence $|\Lambda'| = |\Gamma^{-1}(\sigma - 1)\Lambda| = |\Gamma^{-1}|$ and $(\sigma - 1)\Lambda' = 1 + o(1)$. For any such Λ',

$$(\sigma - 1)\wp(\Lambda') = \wp(1 + o(1)) + p\phi(\Lambda', 1 + o(1)) = o(1) + O(p\Lambda'^{p-1}) = o(1)$$

from definition 5 because $|\Lambda'^{p-1}| < |p^{-1}|$ by (17). From (13) we see there is a $\lambda \varepsilon k$ such that $|\wp(\Lambda') - \lambda| = |\Gamma^{-1}(\sigma - 1)\wp(\Lambda')| < |\Gamma^{-1}| = |\Lambda'|$. This proves (18).

Finally assume K/k is ramified and condition (15) holds. Then $|\Lambda'| \not\leqq |k^{\cdot}|$ in the previous discussion so $|\Gamma| \not\leqq |k^{\cdot}|$. Conversely, if K/k is ramified and $|\Gamma| \not\leqq |k^{\cdot}|$ then Γ^{-1} is an orthogenerator and $(\sigma - 1)\delta\Gamma^{-1} = 1 + o(1)$ for some δ in k.

4. Maximally complete fields. Assume that k is maximally complete [2.4] under its valuation $|\ |$, that K/k is cyclic of degree p with generating automorphism σ, that $|\sigma\mathrm{B}| = |\mathrm{B}|$ for all $\mathrm{B} \varepsilon K$, and that $0 \leqq |p| < 1$. Then ef necessarily equals p, K is also maximally complete, and every pseudo-convergent sequence of elements of K has a pseudo-limit in K. For everything concerning maximally complete fields we follow the definitions of Kaplansky [2] and Schilling [4] except that we use the $|\ |$-notation in place of their v-notation for valuations.

THEOREM 3. *If* K/k *is ramified and the assumptions of Section* 4 *hold*

then $|\Gamma|\varepsilon|k^{\cdot}|$ *if and only if* k *contains a primitive* p*-th root of unity and* $K=k(\beta^{1/p})$ *for a* $\beta\varepsilon k$ *with* $|\beta|\notin|k^{\cdot p}|$.

Proof. If $K=k(\beta^{1/p})$ with $|\beta|\notin|k^{\cdot p}|$ and k contains primitive p-th roots of unity then K/k is ramified and $\beta^{1/p}$ is an orthogenerator. Since $\sigma(\beta^{1/p})=\beta^{1/p}\zeta$ for some primitive p-th root of unity ζ, we have $|\Gamma|=|\zeta-1|\varepsilon|k^{\cdot}|$ from Proposition 2.

Let K/k be ramified and $|\Gamma|\varepsilon|k^{\cdot}|$. By Theorems 1 and 2 $|\Gamma^{p-1}|=|p|$ so k must have characteristic zero. Express Γ by a canonical orthobasis. The values of all nonzero terms are different so we get

$$\Gamma=\gamma+o(\gamma),\quad \gamma\varepsilon k. \tag{19}$$

Then $S\Gamma=S\gamma+S(o(\Gamma))=p\gamma+o(S\Gamma)$ by (11). Hence $S(\Gamma)(1+o(1))=p\gamma$ and $S\Gamma=p\gamma(1+o(1))$. From (19) we see also that $\sigma^i\Gamma=\gamma+o(\gamma)$ for all i, so $N\Gamma=\gamma^p(1+o(1))$. By (14) $\gamma^{p-1}=-p(1+o(1))$. Since k is maximally complete and $|p-1|=1$, $1+o(1)$ is always a $(p-1)$-th power in k [4, p. 61]. So $-p$ is a $(p-1)$-th power in k and since k contains a subfield isomorphic to the p-adic rationals it necessarily [7] contains a primitive p-th root of unity.

Thus $K=k(\mathrm{A})$ with $\mathrm{A}^p=\alpha\varepsilon k$. Since we are assuming $e=p$ and $f=1$ it is easy to see that we may assume that either $|\alpha|\notin|k^{\cdot p}|$ or $\alpha=1+o(1)$. But if $\alpha=1+o(1)$ then K contains an element M satisfying (16), which by Theorem 2 contradicts the fact that $|\Gamma^{p-1}|=|p|$. Namely, let $\mathrm{A}^p=1+o(1)\,\varepsilon k$. Then $\mathrm{A}=1+\Theta$ with $|\Theta|<1$ and $S\mathrm{A}=0$ so $S(p^{-1}\Theta)=-1$ and $\mathrm{M}=p^{-1}\Theta$ satisfies (16).

THEOREM 4. *If the assumptions of Section* 4 *hold and* $|\Gamma^{p-1}|>|p|$ *then* K *is generated by a root* Λ *of an equation* (1) *with* $|\Lambda|=|\Gamma^{-1}|$ *and* $(\sigma-1)\Lambda=1+o(1)$.

Proof. Consider the elements Λ of K which satisfy

$$|\Lambda|=|\Gamma^{-1}| \tag{20}$$

and (15). From Theorem 2 it follows that there are such elements. If Λ, Λ' are two of these we put $\Lambda\prec\Lambda'$ whenever

$$|\Lambda'-\Lambda|=|\Gamma^{-1}(\sigma-1)\wp\Lambda| \tag{21}$$

and

$$|(\sigma-1)\wp\Lambda'|<|(\sigma-1)\wp\Lambda|. \tag{22}$$

Let $\mathfrak{P}$ be the family of all well-ordered sets of elements of the type just

described. $\mathfrak{P}$ exists by virtue of Zermelo's *Aussonderungsaxiom,* and is non-empty. For W, $W' \varepsilon \mathfrak{P}$ we shall say W is less than W' when W is an initial segment of W'. Then $\mathfrak{P}$ is partially ordered and every linearly ordered subset of $\mathfrak{P}$ has an upper bound in $\mathfrak{P}$. By the Lemma of Zorn $\mathfrak{P}$ contains a maximal element M.

CONTENTION 1. *M has a last element.*

Proof. From (21) and (22) follows

$$(23) \qquad |\Lambda'' - \Lambda'| < |\Lambda' - \Lambda| \text{ whenever } \Lambda \prec \Lambda' \prec \Lambda'' \text{ and } \Lambda, \Lambda', \Lambda'' \varepsilon M.$$

So if M has no last element it is a pseudo-convergent sequence. Assume that this is so and let Λ^* be a pseudo-limit of this sequence. Then

$$(24) \qquad |\Lambda^* - \Lambda| = |\Lambda' - \Lambda| \text{ whenever } \Lambda \prec \Lambda' \text{ and } \Lambda, \Lambda' \varepsilon M.$$

We shall show that the structure (M, Λ^*) obtained by adjoining Λ^* to M as last term is in $\mathfrak{P}$, contradicting the maximality of M. By (20), (15), and (18) $(\sigma - 1)\wp(\Lambda) = o(1)$ so by (21) $|\Lambda' - \Lambda| < |\Gamma^{-1}|$ and $|\Lambda^*| = |\Gamma^{-1}|$. It remains to check (15) and (22).

Let Λ be any element of M and $\Theta = \Lambda^* - \Lambda$. We have just seen that $\Theta = o(\Gamma^{-1})$. By (12) $(\sigma - 1)\Theta = o(1)$ so Λ^* satisfies (15). Now if $(\sigma - 1)\Theta = \Delta$ then

$$(\sigma - 1)\wp(\Theta) = \wp(\Theta + \Delta) - \wp(\Theta) = \wp(\Delta) + p\phi(\Theta, \Delta) = -\Delta + o(\Delta).$$

We see that (12) implies

$$(25) \qquad (\sigma - 1)\wp(\Theta) = -(\sigma - 1)\Theta + o(\Gamma\Theta).$$

This being so, $|\Lambda| = |\Gamma^{-1}|$ and $|\Gamma^{p-1}| > |p|$ imply

$$(26) \qquad (\sigma - 1)\wp(\Lambda + \Theta) = (\sigma - 1)\wp(\Lambda) - (\sigma - 1)\Theta + o(\Gamma\Theta).$$

From (21) we see that $|(\sigma - 1)\wp(\Lambda)| = |\Gamma\Theta|$. So from (26)

$$(\sigma - 1)\wp(\Lambda^*) = O(\Gamma\Theta) = O((\sigma - 1)\wp\Lambda).$$

But this holds for every $\Lambda \varepsilon M$ so Λ^* satisfies (22). Thus $(M, \Lambda^*) \varepsilon \mathfrak{P}$ and Contention 1 follows.

CONTENTION 2. *If Λ' is the last element of M then $(\sigma - 1)\wp(\Lambda') = 0$. Hence Λ' generates K/k and satisfies an equation* (1).

Proof. Let $(\sigma - 1)\wp(\Lambda') = \Delta$. From Theorem 2 $\Delta = o(1)$. If $\Delta \neq 0$ let Θ be an element of K with $|\Theta| = |\Gamma^{-1}\Delta|$ and $(\sigma - 1)\Theta = \Delta$. Its

existence is guaranteed by (13). Let $\Lambda^* = \Lambda' + \Theta$. We shall show that $(M, \Lambda^*) \in \mathfrak{P}$ giving a contradiction as before. Everything except the analogue of (22) is clear. For this we apply (26).

$$(\sigma - 1)\wp(\Lambda' + \Theta) = (\sigma - 1)\wp(\Lambda') - (\sigma - 1)\Theta + o(\Gamma\Theta) = o(\Delta).$$

This proves Contention 2 and Theorem 4.

THEOREM 5. *Let K/k satisfy the assumptions of Section 4. If K/k is ramified and $|\Gamma^{p-1}| > |p|$ then $K = k(\Lambda)$ where Λ is an orthogenerator satisfying an equation* (1) *with*

$$\lambda \notin |k^{\cdot p}|, \quad |\lambda| > 1. \tag{27}$$

If $\bar{K}/\bar{k}$ is inseparable and $|\Gamma^{p-1}| > |p|$ then $K = k(\Lambda)$ where Λ is an orthogenerator satisfying an equation (1) *with*

$$\lambda = \alpha\beta^p, \quad |\alpha| = 1, \quad \bar{\alpha} \notin \bar{k}^{\cdot p}, \quad \beta \in k, \quad |\beta| > 1. \tag{28}$$

If K/k is separable then $K = k(\Lambda)$ where Λ is an orthogenerator satisfying an equation (1) *with*

$$|\lambda| = 1, \qquad \bar{\lambda} \notin \wp(\bar{k}^+). \tag{29}$$

Conversely, let k satisfy the assumptions of Section 4. Then every polynomial (1) *either splits completely or has a zero Λ which generates a cyclic extension of degree p and satisfies* (15) *for some generating automorphism σ. If $K = k(\Lambda)$ satisfies the assumptions of Section 4 and λ satisfies* (27), (28), *or* (29) *then K/k is ramified, $\bar{K}/\bar{k}$ is inseparable, or $\bar{K}/\bar{k}$ is separable, respectively, and $|\Gamma^{p-1}| > |p|$.*

Proof. Let $|\Gamma^{p-1}| > |p|$. If K/k is ramified then by Theorems 2 and 4 we have $K = k(\Lambda)$ with $|\Lambda| = |\Gamma^{-1}| \notin |k^{\cdot}|$ so $|\lambda| = |\Lambda^p| \notin |k^{\cdot p}|$. If $\bar{K}/\bar{k}$ is inseparable then $|\Gamma| < 1$ and $K = k(\Lambda)$ with $|\Lambda| = |\Gamma^{-1}| > 1$ so $|\lambda| = |\Lambda^p| \in |k^{\cdot p}|$. Let $\lambda = \alpha\beta^p$ with $|\alpha| = 1$, $|\beta| > 1$, $\alpha, \beta \in k$. Then $|\Lambda\beta^{-1}| = 1$ and $(\Lambda\beta^{-1})^p = \alpha + o(1)$. Suppose $\bar{\alpha} = \bar{\gamma}^p$ with $\gamma \in k$. Then $\Lambda\beta^{-1} = \gamma + o(1)$, $\Lambda = \beta\gamma + o(\Lambda)$, and $(\sigma - 1)\Lambda = o(1)$, which is not true, so $\bar{\alpha} \notin \bar{k}^{\cdot p}$. The case $\bar{K}/\bar{k}$ separable is easily proved.

The proof of Proposition 17 of [7] is easily generalized to prove the statement concerning the behavior of equations (1). If $K = k(\Lambda)$ and λ satisfies (27) then K/k is ramified with Λ an orthogenerator. If λ satisfies (28) then the residue class of $\Lambda\beta^{-1}$ generates an inseparable extension of $\bar{k}$ of degree p. The case when λ satisfies (29) is easily proved. $|\Gamma^{p-1}| > |p|$ was established in Theorem 2.

Warning: There exist irreducible equations (1) which do not satisfy

(27), (28), or (29). If $|\lambda^{p-1}| > |p^{-p}|$ they sometimes generate cyclic extensions with $|\Gamma^{p-1}| > |p|$ but the value of Γ cannot be read off directly from the value of λ.

Definition 6. A field generated by a root of an equation (1) is called an *Artin-Schreier extension.*

PROPOSITION 3. *If $\bar{k}$ has no inseparable extensions then every cyclic subfield of a composite of Artin-Schreier extensions of k is an Artin-Schreier extension.*

Proof. By Theorem 4 we can assume k contains primitive p-th roots of unity. Then the Artin-Schreier extensions are those obtained by adjoining a p-th root of an element of value 1, by the proof of Theorem 3.

We do not know whether the assumption on $\bar{k}$ can be omitted.

5. Explicit reciprocity law. Let k satisfy the assumptions of Section 4. Let C be the algebraic closure of k and G the Galois group of C/k [1]. Let I be the set of all $\lambda \varepsilon k$ such that $|\lambda^{p-1}p^p| < 1$.

If $\lambda \varepsilon$ I and $\sigma \varepsilon G$ choose a $\Lambda \varepsilon C$ such that $\wp(\Lambda) = \lambda$ and denote by $\lambda^*\sigma$ the residue class modulo p of an integer ν such that $(\sigma - 1)\Lambda = \nu + o(1)$. ν exists by Theorem 5. It is easy to verify that $\lambda^*\sigma$ does not depend upon the choice of Λ and that $\lambda^*\sigma\tau = \lambda^*\sigma + \lambda^*\tau$ if $\sigma, \tau \varepsilon G$. Hence λ^* is a character of G of period p.

The mapping $\lambda \to \lambda^*$ is not in general a homomorphism (as it is when k has characteristic p). Namely, if $|p|$ is small enough it can happen that both λ and $p\lambda$ lead to equations of type (28). Then $(p\lambda)^* \neq 0$ but $p(\lambda^*) = 0$. We shall now develop a sufficient condition that $(\lambda + \mu)^* = \lambda^* + \mu^*$.

PROPOSITION 4. *If $\Lambda \varepsilon C$, $\lambda \varepsilon$ I, and $\wp(\Lambda) = \lambda + o(1)$ then $k(\Lambda)$ contains a Λ_0 with $\Lambda_0 = \Lambda + o(1)$ and $\wp(\Lambda_0) = \lambda$.*

Proof. If $\Lambda' \varepsilon k(\Lambda)$ and $\wp(\Lambda') = \lambda + o(1)$ let $\Delta' = \wp(\Lambda') - \lambda$ and define $\Lambda'' = \Lambda' + \Delta'$. Then $\wp(\Lambda'') = \wp(\Lambda') + \wp(\Delta') + p\phi(\Lambda', \Delta')$. By definition of I, $|p\Lambda'^{p-1}| < 1$; so $\wp(\Lambda'') = \lambda + o(\Delta')$. Our proposition follows by applying the Lemma of Zorn just as in the proof of Theorem 4. Of course, the Lemma of Zorn can be avoided when the valuation is discrete and of rank 1.

COROLLARY. *If $|\lambda| < 1$ then $\lambda^* = 0$.*

PROPOSITION 5. *If* $|\mu| \leqq |\lambda|$, $\lambda \varepsilon \mathfrak{l}$, *and* $|\lambda^{p-1}\mu p^p| < 1$ *then* $\mu^* + \lambda^* = (\mu + \lambda)^*$.

Proof. Let $\wp(\Lambda) = \lambda$, $\wp(\mathrm{M}) = \mu$. Let $\mathrm{N} = \Lambda + \mathrm{M}$. Then

$$\wp(\mathrm{N}) = \lambda + \mu + p\phi(\Lambda, \mathrm{M}) = \lambda + \mu + o(1).$$

By Proposition 4 there is an N_0 in $k(\mathrm{N})$ such that $\mathrm{N}_0 = \mathrm{N} + o(1)$ and $\wp \mathrm{N}_0 = \lambda + \mu$. If $\sigma \varepsilon G$ then

$$(\sigma - 1)\mathrm{N}_0 = (\sigma - 1)\mathrm{N} + o(1) = \phi^*\sigma + \mu^*\sigma + o(1)$$

which was to be proved.

In particular, if $|\mu| = 1$ and $\lambda \varepsilon \mathfrak{l}$ then we have $\mu^* + \lambda^* = (\mu + \lambda)^*$.

For each $\lambda \varepsilon \mathfrak{l}$ we can form the 2-cocycle (λ^*, α, k) according to [7] or the algebra $(\alpha, \lambda]$ according to [10], which amounts to the same thing. Consider the symbol $(\alpha, \lambda]$ as denoting the corresponding 2-cohomology class as in [7]. Using Theorem 1 of [7] and Proposition 5 of this paper we obtain the following rules:

1. *For all* $\lambda \varepsilon \mathfrak{l}$, $(\lambda, -\lambda] = 1$.
2. *If* $\lambda, \mu \varepsilon \mathfrak{l}$, $|\mu| \leqq |\lambda|$, *and* $|\lambda^{p-1}\mu p^p| < 1$ *then* $(\alpha, \lambda](\alpha, \mu] = (\alpha, \lambda + \mu]$. *This is always true if* $\lambda \varepsilon \mathfrak{l}$ *and* $|\mu| = 1$.
3. $(\alpha, \lambda](\beta, \lambda] = (\alpha\beta, \lambda]$.

Rule 1 follows from the fact that $-\lambda$ is the norm of $-\Lambda$ if $\wp(\Lambda) = \lambda$.

We now assume that k is a regular field of generalized local class field theory [7].

Let $\lambda = \lambda'\pi^{-c} \varepsilon \mathfrak{l}$, $|\lambda'| = 1$, and $p \nmid c$. We wish to compute the invariant of the algebra $(\beta, \lambda]$. If we could do this for every $\beta \varepsilon k^\cdot$ we would have a complete explicit reciprocity law. As it is, we are able to compute the invariant of $(\beta, \lambda]$ for a special class of elements β which form a basis for $k^\cdot$ modulo $N(K/k)$.

Let α be an element of k such that $K(\mathrm{A})/k$ is unramified if $\wp(\mathrm{A}) = \alpha$ and such that the invariant $\rho(\pi, \alpha]$ (see [7]) is $[1/p] \pmod 1$. Then, since $|\alpha - \lambda| = |\pi^{-c}|$, $\rho(\alpha - \lambda, \alpha] = [-c/p] \pmod 1$. By rule 1,

$$(\alpha - \lambda, -(\alpha - \lambda](\alpha - \lambda, \alpha] = (\alpha - \lambda, \alpha].$$

By rule 2,

$$(\alpha - \lambda, -(\alpha - \lambda)](\alpha - \lambda, \alpha] = (\alpha - \lambda, \lambda].$$

Hence $\rho(\alpha - \lambda, \lambda] = \rho(\alpha - \lambda, \alpha] = [-c/p] \pmod 1$. If one prefers a β

of value 1 instead of $|\pi^{-c}|$ then $\rho(1-\alpha\lambda^{-1},\lambda]=[-c/p] \pmod 1$ since $-\lambda^{-1}$ is a norm. Thus we obtain the law

$$\rho(1-\alpha\lambda^{-1},\lambda]=[-c/p] \pmod 1. \tag{30}$$

We are now in a position to prove the explicit reciprocity law for the field of primitive p^2-th roots of unity over the p-adic completion of the rational field. If this problem were solved for the field of p^n-th roots of unity it would make possible an elegant proof of Artin's global reciprocity law, for it is well known that this law can be proved from the global index theorems and the global reciprocity law for cyclotomic fields.

Appendix.

SCHOLIUM 1. *It is not possible without assumptions on k in addition to those of Sections* 1, 2, 3 *to prove that every cyclic K/k with* $|\Gamma^{p-1}|>|p|$ *is generated by a root of an equation* (1).

Proof. Let r denote the rational field, $|\ |$ the p-adic valuation, and C the cyclic subfield of degree p of the field of primitive p^2-th roots of unity. Then $e=p$. If r_p denotes the p-adic rationals then Cr_p/r_p is generated by a root of $x^p-x-p^{-1}=0$ (see [8]) so $|\Gamma^{p-1}|>|p|$. But for p odd C/r is cyclic of odd degree so any defining equation must split into real linear factors in C. By Descartes' rule of signs an equation $x^p-x-a=0$, for a real, can have at most three real roots. Hence C/r is not generated by a root of any such equation for $p>3$.

Unfortunately this method does not disprove the conjecture: if $|\Gamma^{p-1}|>|p|$ and k contains primitive $(p-1)$-th roots of unity then K is generated by a root of an equation (1). Although this seems very unlikely we have not been able to find a counterexample.

SCHOLIUM 2. *Without assumptions in addition to those of Sections* 1, 2, 3 *it is not possible to prove that every irreducible equation* (1) *generates a cyclic extension for $p>2$.*

Proof. Consider x^p-x-1 over the rational field. Being irreducible modulo p is it irreducible. For $p>3$ Descartes' rule of signs shows that its splitting field has even degree and this is easy to check for $p=3$ also.

SCHOLIUM 3. *If k is not complete then K/k cyclic, $e=p$, $|\Gamma^{p-1}|=|p|$ can happen without primitive p-th roots of unity being contained in k.*

Proof. $r(\sqrt{6})$ does not contain primitive cube roots of unity (being

real) but $r_3(\sqrt{6})$ does. So by Wang's Theorem [5] it is easy to see that there is a cyclic cubic extension of $r(\sqrt{6})$ which is a counterexample.

SCHOLIUM 4. *If $\bar{k}$ has an inseparable extension there may exist a K/k which is cyclic of degree p, does not contain primitive p-th roots of unity, and contains no Λ with $(\sigma-1)\Lambda=1+o(1)$.*

Proof. Let p be any odd prime. Let k_0 be the completion of the field of all rational functions with rational coefficients of a transcendental element t, under the valuation which defines the value of a polynomial in t to be the maximum p-adic value of its coefficients. Let $k=k_0(((1-t^{p-1})p)^{1/(p-1)})$. Let K be the splitting field over k of the polynomial x^p-px-t.

INDIANA UNIVERSITY AND INSTITUTE FOR ADVANCED STUDY.

REFERENCES.

[1] E. Artin, *Algebraic Numbers and Algebraic Functions. I.* Mimeographed lecture notes, Princeton University, New York University, 1951.

[2] I. Kaplansky, "Maximal fields with valuation," *Duke Mathematical Journal*, vol. 9 (1942), pp. 303-321.

[3] O. Ore, "Abriss einer arithmetischen Theorie der Galoisschen Körper," Parts 1 and 2, *Mathematische Annalen*, vol. 100 (1928), pp. 650-673, and vol. 102 (1930), pp. 283-304.

[4] O. F. G. Schilling, *The Theory of Valuations*, Mathematical Surveys, no. 4, American Mathematical Society, New York, 1950.

[5] Shianghaw Wang, "On Grunwald's theorem," *Annals of Mathematics*, vol. 51 (1950), pp. 471-484.

[6] G. Whaples, "Existence of generalized local class fields," *Proceedings of the National Academy of Sciences*, vol. 39 (1953), pp. 1100-1103.

[7] ———, "Generalized local class field theory, I and II," *Duke Mathematical Journal*, vol. 19 (1952), pp. 505-517, and vol. 21 (1954), pp. 247-255.

[8] ———, "Generalized local class field theory, IV, Cardinalities," *ibid.*, vol. 21 (1954), pp. 583-586.

[9] E. Witt, "Der Existenzsatz für abelsche Funktionenkörper," *Journal für die reine und angewandte Mathematik*, vol. 173 (1935), pp. 43-51.

[10] ———, "Schiefkörper über diskret bewerteten Körpern," *ibid.*, vol. 176 (1936), pp. 153-156.

REMARK ON AN APPLICATION OF PSEUDOANALYTIC FUNCTIONS.*[1]

By LIPMAN BERS.

1. Introduction. In this note we show how the theory of pseudoanalytic functions as formulated in [5] yields very precise information on the behaviour of solutions of a linear ellipitic equation

$$(1.1) \qquad L\phi = a_{11}\phi_{xx} + 2a_{12}\phi_{xy} + a_{22}\phi_{yy} + a_1\phi_x + a_2\phi_y + a_0\phi = 0$$

near a regular or isolated singular point. The theorems stated below (Section 2) contain and are stronger than the previous results on local behaviour of solutions of (1.1) due to the author [1, 2, 4], Vekua [16] and Hartman and Wintner [10, 12]. These theorems refer to single-valued solutions. But the same method would give analogous results for finitely-many-valued solutions and for solutions with finitely-many-valued gradients.

Without loss of generality we consider equations and solutions defined near the origin of the z-plane ($z = x + iy$) and assume that

$$(1.2) \qquad a_{11} = a_{22} = 1, \qquad a_{12} = 0 \text{ at } z = 0.$$

Concerning the smoothness of the coefficients in (1.1) we make either of the following assumptions.

HYPOTHESIS α. *The leading coefficients $a_{ij}(x, y)$ of* (1.1) *satisy a uniform Hölder condition and condition* (1.2). *The coefficients $a_i(x, y)$ are measurable functions which belong to the space L_p for some $p > 2$.*

HYPOTHESIS β. *All coefficients of* (1.1) *satisfy a uniform Hölder condition and* (1.2) *holds.*

Under Hypothesis β we require that solutions of (1.1) be of class C^2. Under Hypothesis α we have no right to expect twice continuously differen-

* Received November 28, 1955.

[1] Work supported by the Office of Ordnance Research, United States Army, under Contract No. DA-30-069-ORD-835. Reproduction in whole or in part is permitted for any purpose of the United States Government.

tiable solutions. We shall say that ϕ is a solution if the derivatives ϕ_x, ϕ_y exist, are continuous and possess generalized L_p derivatives $\phi_{xx} = (\phi_x)_x$, $\phi_{xy} = (\phi_x)_y = (\phi_y)_x$ and $\phi_{yy} = (\phi_y)_y$ which satisfy (1.1) almost everywhere. The equality of the mixed derivatives is a consequence of the definition of generalized derivatives.

In Section 3 we shall show, using an inequality of Calderón and Zygmund [7], that equation (1.1) can be reduced to an equation involving only second derivatives (Theorem I), to a system of two first order equations (Theorem II), and to the equations characterizing pseudoanalytic functions (Theorem III). This will yield the desired description of the local behaviour of solutions (Theorems A and B of Section 2).

Hypothesis α is essentially sharp, since our results are certainly not true for equations with continuous but not Hölder continuous coefficients (cf. Hartman and Wintner [12]). Concerning the local behaviour of solutions of elliptic equations with discontinuous coefficients we refer to a recent paper by Nirenberg and the author [6].

For the convenience of the reader we recall the definition of generalized derivatives (Friedrichs [8], Sobolev [15]). Let $f(x,y)$, $g(x,y)$, $h(x,y)$ be measurable functions defined in a domain D. If $|f|^p$, $|g|^p$, $|h|^p$ are locally summable ($p \geqq 1$), the statement

$$f_x = g,\ f_y = h \text{ in the } L_p\text{-sense}$$

means that

(a) in every compact subset D_0 of D there exists a sequence of continuously differentiable functions $f^{(n)}$ such that

$$\int\int_{D_0} \{|f^{(n)} - f|^p + |f_x^{(n)} - g|^p + |f_y^{(n)} - h|^p\} dxdy \to 0,$$

(b) for every continuously differentiable function ω which vanishes identically near the boundary of D

$$\int\int_D f\omega_x\, dxdy = -\int\int_D g\omega\, dxdy, \quad \int\int_D f\omega_y\, dxdy = -\int\int_D h\omega\, dxdy,$$

and

(c) for almost every y and for almost every x

$$\int_\alpha^\beta g(\xi, y)d\xi = f(\beta, y) - f(\alpha, y), \qquad \int_\gamma^\delta h(x, \eta)d\eta = f(x, \delta) - f(x, \gamma)$$

for almost all $\alpha, \beta, \gamma, \delta$.

Each of the three conditions (a), (b), (c) implies the other two. It is known that f is Hölder continuous if f_x and f_y exist in the L_p-sense for some $p > 2$.

2. Statement of results. Let $\phi(z) = \phi(x, y)$ be a solution of (1.1) defined in a neighborhood of the origin. Assuming Hypothesis α we say that ϕ has a zero of integral order $n > 0$ at the origin if for some complex $\alpha \neq 0$

$$\phi \sim \mathrm{Re}(\alpha z^n), \qquad \phi_x - i\phi_y \sim n\alpha z^{n-1}. \tag{2.1}$$

In the case of Hypothesis β we also require that

$$\phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy} \sim n(n-1)\alpha z^{n-2}. \tag{2.2}$$

Here and hereafter all asymptotic relations refer to $z \to 0$.

THEOREM A. *Under Hypotheses α or β a solution of* (1.1) *which vanishes at the origin without vanishing identically has at the origin a zero of integral order.*

The following consequence of Theorem A is of interest for differential geometry in the large (cf. Hartman and Wintner [11] and the references given there).

COROLLARY. *Let the function $\Omega(x, y, \phi, p, q, r, s, t)$ be a Hölder continuously differentiable function of its eight variables, defined for sufficiently small values of the variables, and satisfying the condition: $4\Omega_r\Omega_t - \Omega_s^2 > 0$ at $x = y = \phi = p = q = r = s = t = 0$. Let $\phi'(x, y)$ and $\phi''(x, y)$ be two twice continuously differntiable functions defined in the neighborhood of the origin and satisfying the differential equation*

$$\Omega(x, y, \phi, \phi_x, \phi_y, \phi_{xx}, \phi_{xy}, \phi_{yy}) = 0. \tag{2.3}$$

Assume that the functions ϕ', ϕ'' vanish at the origin together with their derivatives of the first and second order, but the difference $\phi = \phi'' - \phi'$ is not identically zero. Then the expression $\phi_{xx}\phi_{yy} - \phi_{xy}^2$ is negative in a deleted neighborhood of the origin.

The conclusion of the Corollary was stated, proved and applied to the uniqueness proof for Minkowski's problem by H. Lewy [13], under the assumption that the function Ω, and hence also every solution of the differential equation (2), is analytic. Recently Hartman and Wintner [9, 10, 11] extended it to the case when Ω is twice continuously differentiable. The

present formulation seems to be essentially sharp. It will be seen from the proof, however, that for special cases (for instance, for quasi-linear equations) weaker conditions are sufficient.

In order to prove the Corollary observe that the difference $\phi = \phi'' - \phi'$ satisfies a linear elliptic partial differential equation of the form (1.1) with

$$a_{11}(x,y) = \int_0^1 \Omega_r[x, y, (1-\lambda)\phi'(x,y) + \lambda\phi''(x,y), \cdots]d\lambda$$

and the other coefficients a_{ij}, a_i defined similarly. We may assume that (1.2) holds since this can be achieved by an an affine transformation. According to a theorem of Nirenberg [14] the second derivatives of the solutions ϕ', ϕ'' satisfy a Hölder condition. Hence the coefficients a_{ij}, a_i satisfy such a condition, and Theorem A implies that $\phi_{xx}\phi_{yy} - \phi_{xy}^2 \sim -n^2(n+1)^2 |\alpha|^2 |z|^{2n-2}$ with $|\alpha| > 0$ and $n > 1$.

We consider next a single-valued solution ϕ of (1.1) defined in a deleted neighborhood of the origin. This solution will be said to have at the origin a logarithmic singularity if (under Hypothesis α) for some real $\alpha \neq 0$

$$\phi \sim \alpha \log|z|, \qquad \phi_x - i\phi_y \sim \alpha/z \tag{2.4}$$

and (under Hypothesis β) also

$$\phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy} \sim -\alpha/z^2. \tag{2.5}$$

We say that the origin is a pole of ϕ if (under Hypothesis α)

$$\phi \sim \mathrm{Re}(\alpha z^{-n}), \qquad \phi_x - i\phi_y \sim -n\alpha z^{-n-1} \tag{2.6}$$

for some complex $\alpha \neq 0$ and integer $n > 0$, and (under Hypothesis β) also

$$\phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy} \sim n(n+1)z^{-n-2}. \tag{2.7}$$

Finally, $z = 0$ will be called an essential singularity of ϕ if (under Hypothesis α)

$$\begin{aligned} \limsup(|z|^N \phi) &= \limsup(-|z|^N \phi) \\ &= \limsup(|z|^N |\phi_x - i\phi_y|) = +\infty \end{aligned} \tag{2.8}$$

and (under Hypothesis β) also

$$\limsup(|z|^N |\phi_{xx} - i\phi_{xy}|) = \limsup(|z|^N |\phi_{yy} + i\phi_{xy}|) = +\infty \tag{2.9}$$

for every $N > 0$. If $a_0 \equiv 0$, we require also that

$$\liminf |\phi_x - i\phi_y - \gamma| = 0 \tag{2.10}$$

for every complex γ.

THEOREM B. *Under Hypotheses α or β let ϕ be a single-valued solution of (1.1) defined in a deleted neighborhood of the origin. Then the singularity of ϕ at $z=0$ is either removable, or logarithmic, or a pole, or essential.*

The theorem implies, in particular, that if ϕ' and ϕ'' are two solutions having at $z=0$ logarithmic singularities, then there exist two real constants λ_1, λ_2 such that $\lambda_1\phi' + \lambda_2\phi''$ is regular at the origin.

3. Reduction to pseudoanalytic functions. This reduction will be accomplished in three steps.

We remark that in the proofs of Lemma 3.1 and Theorem I given below the Hölder continuity of the leading coefficients and the fact that the number of independent variables is two are not used, and the conclusions would remain true, with obvious modifications, in a more general case, say for continuous a_{ik}.

LEMMA 3.1. *Under Hypothesis α there exists a solution ϕ of equation (1.1) which is defined in a neighborhood of the origin and satisfies the conditions*

$$\phi = c_0,\ \phi_x = c_1,\ \phi_y = c_2 \text{ at } z = 0, \tag{3.1}$$

where c_0, c_1, c_2 are given numbers.

The proof is a variant of the classical Korn argument based on the results of Calderón and Zygmund [7]. Let $\boldsymbol{B}$ denote the Banach space of real functions $\phi(z)$ defined for $|z| \leqq R$ and possessing continuous first and generalized second derivatives, for which the norm

$$\|\phi\| = \max(|\phi|, |\phi_x|, |\phi_y|) + \{\iint_{|z|<R} (|\phi_{xx}| + |\phi_{xy}| + |\phi_{yy}|)^p\, dxdy\}^{1/p}$$

is finite. Define the mapping $\boldsymbol{T}$ of $\boldsymbol{B}$ into itself by defining $\chi = \boldsymbol{T}\phi$ to be the function

$$\chi(z) = \tfrac{1}{2\pi} \iint_{|\zeta|<R} [(\Delta - \boldsymbol{L})\phi(\zeta)] \log|\zeta - z|\, d\xi d\eta.$$

Using the inequalities of Hölder and of Calderón-Zygmund it is easy to show that $\boldsymbol{T}$ is bounded, that $\Delta\boldsymbol{T} = \Delta - \boldsymbol{L}$, and that $\|\boldsymbol{T}\| \to 0$ for $R \to 0$. If $h(z)$ is a harmonic function with $\|h\| < +\infty$, and R is so small that $\|\boldsymbol{T}\| \leqq \theta < 1$, then the equation

$$\phi = \boldsymbol{T}\phi + h \tag{3.2}$$

has a solution which satisfies equation (1.1) and the inequality

$$\| \phi - h \| \leqq \theta(1-\theta)^{-1} \| h \|.$$

We solve equation (3.2) for a sufficiently small R and for $h = 1, x, y,$ and obtain three solutions ϕ_0, ϕ_1, ϕ_2 of (1.1) for which the values of

$$| \phi_0 - 1 |, | \phi_{0,x} |, | \phi_{0,y} |, | \phi_1 |, | \phi_{1,x} - 1 |, | \phi_{1,y} |, | \phi_2 |, | \phi_{2,x} | | \phi_{2,y} - 1 |$$

are very small. A properly chosen linear combination of these solutions satisfies (3.1).

LEMMA 2.2. *Under Hypothesis β there exists a solution ϕ of equation (1.1) which is defined in the neighborhood of the origin and satisfies the conditions*

$$(2.3) \quad \phi = c_0,\ \phi_x = c_1,\ \phi_y = c_2,\ \phi_{xx} = -\phi_{yy} = c_{11},\ \phi_{xy} = c_{12} \text{ at } z = 0,$$

where c_0, c_1, c_2, c_{11}, c_{12} are given numbers.

This (known) result is proved by considering $\boldsymbol{T}$ as an operator in the Banach space of functions ϕ defined in $|z| \leqq R$ and having second derivatives satisfying a uniform Hölder condition with exponent $\epsilon < 1$, ϵ being a Hölder exponent for the coefficients of (1.1). The norm in this space is

$$\| \phi \| = \max(|\phi|, R|\phi_x|, R|\phi_y|, R^2|\phi_{xx}|, R^2|\phi_{xy}|, R^2|\phi_{yy}|) + R^{2+\epsilon}K,$$

where K is the smallest Hölder constant for ϕ_{xx}, ϕ_{xy}, ϕ_{yy} belonging to the exponent ϵ. The desired function is obtained by solving equation (3.2) for a small R and for $h = 1, x, y, x^2 - y^2, xy$.

THEOREM 1. *Under Hypothesis α there exist, in a neighborhood of the origin, three functions ϕ_0, ξ, η such that*

$$\boldsymbol{L}\phi_0 = \boldsymbol{L}\xi = \boldsymbol{L}\eta = 0,$$

$$(3.4) \quad \phi_0 = 1,\ \phi_{0,x} = \phi_{0,y} = 0 \text{ at } z = 0,$$

$$(3.5) \quad \xi = \eta = \xi_y = \eta_x = 0,\ \xi_x = \eta_y = 1 \text{ at } z = 0.$$

A function $\phi(x, y)$ defined in a neighborhood of the origin is a solution of (1.1) if and only if the function $\Phi = \phi/\phi_0$, considered as a function of (ξ, η) is a (twice continuously differentiable) solution of the equation

$$(3.6) \quad A_{11}\Phi_{\xi\xi} + 2A_{12}\Phi_{\xi\eta} + A_{22}\Phi_{\eta\eta} = 0$$

where the A_{ij} are certain functions satisfying a Hölder condition and the condition

$$A_{11} = A_{12} = 1,\ A_{12} = 0 \text{ at } \zeta = \xi + i\eta = 0. \tag{3.7}$$

Under Hypothesis β the same conclusion holds and the functions ϕ, ξ, η may be chosen so that

$$\phi_{0,xx} = \phi_{0,xy} = \phi_{0,yy} = \xi_{xx} = \xi_{xy} = \xi_{yy} = \eta_{xx} = \eta_{xy} = \eta_{yy} = 0 \text{ at } z = 0. \tag{3.8}$$

Proof. The existence of solutions of (1.1) satisfying conditions (3.4), (3.5), (3.8) follows from Lemmas 3.1 and 3.2. The equivalence of equations (1.1) and (3.6) is proved by a direct computation. In the case of Hypothesis α the legitimacy of this computation may be established by a simple argument.

The Hölder continuity of the functions A_{ij} follows from the Hölder continuity of the first derivatives of solutions of (1.1) under Hypothesis α (cf. Section 1).

In order to complete the proof we must show that every solution of (3.6) which has continuous first derivatives and generalized second derivatives in L_p $(p \geqq 2)$ also has continuous second derivatives. This, however, is an immediate consequence of two statements: the Dirichlet problem for an elliptic equation (3.6) with Hölder continuous coefficients has a unique twice continuously differentiable solution; every solution of (3.6) with continuous first and generalized second L_p-derivatives, $p \geqq 2$, obeys the maximum principle. The first statement is classical; the second has been established recently (Bers and Nirenberg [6]).

THEOREM II. *Let there be given an elliptic equation*

$$a_{11}\phi_{xx} + 2a_{12}\phi_{xy} + a_{22}\phi_{yy} = 0 \tag{3.9}$$

satisfying Hypothesis β. There exist four Hölder continuous functions $b_{11}, \cdots, b_{22}$ defined in a neighborhood of the origin, satisfying the condition

$$b_{11} = b_{22} = 0,\ b_{12} = -b_{21} = 1 \text{ at } z = 0 \tag{3.10}$$

and such that in a neighborhood of the origin equation (3.9) is equivalent to the elliptic system

$$\phi_x = b_{11}\psi_x + b_{12}\psi_y, \qquad \phi_y = b_{21}\psi_x + b_{22}\psi_y. \tag{3.11}$$

This means that to every solution ϕ of (3.9) there exists a (not necessarily single-valued) function ψ satisfying (3.11), and that whenever ϕ and

ψ have continuous derivatives and satisfy (3.11), ϕ is a twice continuously differentiable solution of (3.9).

Addition to Theorem II. *Let the coefficients of* (3.9) *be Hölder continuous and satisfy the ellipticity condition* $(a_{11}a_{22} - a_{12}^2 > 0)$ *in a domain D. Then there exist Hölder continuous functions* b_{ij} *defined in the whole domain D, satisfying* (3.10), *and such that system* (3.11) *is elliptic* $(-4b_{12}b_{21} - b_{11}^2 - b_{22}^2 + 2b_{11}b_{22} > 0)$ *and equivalent to* (3.9).

Proof. Consider the elliptic system

$$\lambda_x = (2a_{12}/a_{22})\mu_x + \mu_y, \qquad \lambda_y = -(a_{11}/a_{22})\mu_x. \tag{3.12}$$

Using the Korn-Lichtenstein method (cf. the proof of Lemmas 2.1, 2.2) or the theorem on univalent solutions of elliptic systems [3] one obtains easily a solution (λ, μ) with Hölder continuous derivatives, defined in the neighborhood of the origin and satisfying the conditions

$$\lambda_x = \mu_y = 0, \quad -\lambda_y = \mu_x = 1 \text{ at } z = 0. \tag{3.13}$$

With these functions form the system

$$\chi_x = \lambda_x\phi_x + \mu_x\phi_y, \qquad \chi_y = \lambda_y\phi_x + \mu_y\phi_y. \tag{3.14}$$

We claim that near $z = 0$ every solution ϕ of (3.9) is a solution of (3.14), that is that

$$I_C = \int_C (\lambda_x\phi_x + \mu_x\phi_y)\,dx + (\lambda_y\phi_x + \mu_y\phi_y)\,dy = 0,$$

where C is a simple closed smooth curve located near the origin and homotopic to zero in the domain of definition of ϕ. This would follow at once from Green's theorem and (3.12), if the functions λ, μ were of class C^2. Under the present circumstances the proof can be accomplished by approximating λ and μ by C^2 functions.

Conversely, if (χ, ϕ) satisfies (3.14), ϕ is a solution of (3.9). This follows from the unique solvability of the Dirichlet problem for (3.9) since solutions of (3.14) obey the maximum principle (cf., for instance, [6]).

Solving system (3.14) algebraically for ϕ_x and ϕ_y and setting $\psi = -\chi$ we obtain the desired system (3.11).

In order to prove the Addition to Theorem II we need a solution of (3.12) defined and satisfying the condition $\mu_x > 0$ in the whole domain D. Such a solution can be obtained by the method used in [3]. We give no details since the in-the-large result will not be used in the sequel.

THEOREM III. *Under the hypothesis of Theorem* II *let* $\zeta = \xi(x, y)$

$+i\eta(x,y)$ be a Hölder continuously differentiable homeomorphism of a neighborhood of the origin, with

$$\xi = \eta = \xi_y = \eta_x = 0,\ \xi_x = \eta_y = 1 \text{ at } z = 0, \tag{3.15}$$

which is conformal with respect to the metric

$$a_{22}dx^2 - 2a_{12}\,dxdy + a_{22}dy^2. \tag{3.16}$$

Set

$$G(\zeta) = -(b_{11}+b_{22})/2 + i[-b_{12}b_{21} - (b_{11}-b_{22})^2/4]^{\frac{1}{2}},$$

$$\Gamma(\zeta) = -a_{12}/a_{11} + i(a_{11}a_{22} - a_{12})^{\frac{1}{2}}/a_{11}.$$

Let ϕ be a solution of (3.9) defined in a (perhaps deleted) neighborhood of the origin, ψ the function connected with ϕ by equations (3.11), and set

$$\omega(\zeta) = \phi + i\psi, \qquad w(\zeta) = \phi + G\psi,$$

$$\Omega(\zeta) = \phi_x - i\phi_y, \qquad W(\zeta) = \phi_x - \Gamma\phi_y.$$

Then ω and Ω are pseudoanalytic functions (of the second kind) with generators $(1, G)$ and $(1, \Gamma)$, respectively, w and W are the corresponding pseudoanalytic functions of the first kind, and

$$\omega \sim w, \tag{3.17}$$

$$\Omega \sim W \sim \dot{w}. \tag{3.18}$$

$$\phi_{xx} - i\phi_{xy} \sim -\phi_{yy} + i\phi_{xy} \sim \dot{W}, \tag{3.19}$$

where $\dot{w}$ is the $(1, G)$ derivative of $w(\zeta)$ and W the $(1, \Gamma)$ derivative of $\dot{W}(\zeta)$.

Proof. The existence of the homeomorphism ζ is a classical result of Lichtenstein. Conformality with respect to (3.10) means, of course that

$$\xi_x^2 + \eta_x^2 = \nu a_{22},\ \xi_x\xi_y + \eta_x\eta_y = -\nu a_{12},\ \xi_y^2 + \eta_y^2 = \nu a_{11},$$

where $\nu > 0$. A direct computation shows that the mapping $z \to \zeta$ takes system (3.11) into the system

$$\phi_\xi = \tau\psi_\xi + \sigma\psi_\eta, \qquad \phi_\eta = -\sigma\psi_\xi + \tau\psi_\eta, \tag{3.20}$$

where

$$2\tau = b_{11} + b_{22},\ \sigma^2 + \tau^2 = b_{11}b_{22} - b_{12}b_{21},\ \sigma > 0.$$

These equations express the $(1, G)$ pseudoanalyticity of ω and w. Since $G(0) = i$ in view of (3.10), relation (3.17) follows.

Next, set $\Phi = \phi_x$, $\Psi = -\phi_y$. By (3.9)

$$\Phi_x = (2a_{12}/a_{11})\Psi_x + (a_{22}/a_{11})\Psi_y, \qquad \Phi_y = -\Psi_y.$$

The mapping $z \to \zeta$ takes this system into the system

$$\Phi_\xi = \tau^* \Psi_\xi + \sigma^* \Phi_\eta, \qquad \Phi_\eta = -\sigma^* \Psi_\xi + \tau^* \Psi_\eta, \tag{3.21}$$

where

$$2\tau^* = 2a_{12}/a_{11}, \quad \sigma^{*2} + \tau^{*2} = a_{22}/a_{11}, \quad \sigma^* > 0.$$

These equations express the $(1, \Gamma)$ pseudoanalyticity of $\Omega = \Phi + i\Psi$ and $W = \Phi + \Gamma\Psi$. Since $\Gamma(0) = i$ in view of (1.2), $\Omega \sim W$.

By the definition of the derivative of a pseudoanalytic function,

$$2\dot{w} = \phi_\xi - i\phi_\eta + (\psi_\xi - i\psi_\eta)G, \text{ and } \phi_\xi + i\phi_\eta + (\psi_\xi + i\psi_\eta)G = 0,$$

by (3.20); so that $\dot{w} \sim \phi_\xi - i\phi_\eta$. But by (3.15), $\Omega = \phi_x - i\phi_y \sim \phi_\xi - i\phi_\eta$, so that (3.18) follows.

Also,

$$2\dot{W} = (\Phi_\xi - i\Phi_\eta) + (\Psi_\xi - i\Psi_\eta)\Gamma, \text{ and } \Phi_\xi + i\Phi_\eta + (\Psi_\xi + i\Psi_\eta)\Gamma = 0,$$

by (3.21); so that $\Phi_\xi - i\Phi_\eta \sim W$. But by (3.18), $\Phi_\xi \sim \Phi_x = \phi_{xx}$ and $\Phi_\eta \sim \Phi_y = -\phi_{xy}$ and by (1.2), (3.9) $\phi_{xx} - i\phi_{xy} \sim -\phi_{yy} - i\phi_{xy}$, so that (3.19) is proved.

4. Proof of Theorems A and B. In view of Theorem I it suffices to prove Theorems A and B for an equation of the form (3.6) with Hölder continuous coefficients satisfying (3.7). (In connection with requirement (2.10) for an essential singularity, note that if $a_0 \equiv 0$, we may set $\phi_0 \equiv 1$). Theorems II and III reduce the assertions to be proved to corresponding statements about pseudoanalytic functions established in [5].

NEW YORK UNIVERSITY.

REFERENCES.

[1] L. Bers, "Partial differential equations and generalized analytic functions," *Proceedings of the National Academy of Sciences*, vol. 36 (1950), pp. 130-136. Second note, *ibid.*, vol. 37 (1951), pp. 42-47.

[2] ———, *Theory of pseudo-analytic functions*, Lecture notes (mimeographed), New York University, 1953.

[3] ———, "Univalent solutions of linear elliptic systems," *Communications on Pure and Applied Mathematics*, vol. 6 (1953), pp. 513-526.

[4] ———, "Function-theoretical properties of solutions of partial differential equations of elliptic type," *Annals of Mathematics*, Study 33 (1954), pp. 69-94.

[5] ———, Local theory of pseudo-analytic functions, *Lectures on Functions of a Complex Variable*, 1955, pp. 213-244.

[6] L. Bers and L. Nirenberg, "On a representation theorem for linear elliptic systems with discontinuous coefficients and its applications," *Convegno Internazionale sulle Equazioni Derivate e Parziali*, Agosto 1954, pp. 111-140.

[7] A. P. Calderón and A. Zygmund, "On the existence of certain singular integrals," *Acta Mathematica*, vol. 88 (1952), pp. 85-139.

[8] K. O. Friedrichs,, "The identity of weak and strong extensions of differential operators," *Transactions of the American Mathematical Society*, vol. 55 (1944), pp. 132-151.

[9] P. Hartman and A. Wintner, "On the third fundamental form of a surface," *American Journal of Mathematics*, vol. 75 (1953), pp. 298-334.

[10] ——— and A. Wintner, "On the local behavior of solutions of non-parabolic partial differential equations," *ibid.*, vol. 75 (1953), pp. 449-476.

[11] ——— and A. Wintner, "Umbilical points and W-surfaces," *ibid.*, vol. 76 (1954), pp. 502-508.

[12] ——— and A. Wintner, "On the local behavior of solutions of non-parabolic partial differential equations. II. The uniqueness of the Green singularity," *ibid.*, vol. 76 (1954), pp. 351-361.

[13] H. Lewy, "On differential geometry in the large, I," *Transactions of the American Mathematical Society*, vol. 43 (1938), pp. 258-270.

[14] L. Nirenberg, "On nonlinear elliptic partial differential equations and Hölder continuity," *Communications on Pure and Applied Mathematics*, vol. 6 (1953), pp. 97-156.

[15] S. L. Sobolev, *Some applications of functional analysis*, Leningrad, 1950 (in Russian).

[16] I. N. Vekua, "Systems of partial differential equations of first order of elliptic type and boundary value problems with applications to the theory of shells," *Matematicheski Sbornik*, vol. 31 [73] (1952), pp. 217-314 (in Russian).

GENERALIZED LAPLACIANS.*

By VICTOR L. SHAPIRO.[1]

1. Introduction. The primary purpose of this paper is to answer the following question:

If $f(x)$ and its first r generalized Laplacians are known in a domain G where $f(x)$ and the first $r-1$ generalized Laplacians are continuous and the r-th generalized Laplacian is integrable, can $f(x)$ be obtained from its r-th generalized Laplacian by the expected integral representation?

The answer is in the affirmative, and we prove it by means of Rudin's result [5] on the first generalized Laplacian and by means of Fourier analysis. In so doing, we also solve a question in the uniqueness of multiple trigonometric series left open both by Cheng [3] and the present author [6] as well as obtain an integral depresentation akin to [5] for continuous functions defined on the torus.

2. Definitions and notation. We shall operate in n-dimensional Euclidean space ($n \geqq 2$) designated by E_n and use the following notation:

$$m = (m_1, \cdots, m_n), x = (x_1, \cdots, x_n), \alpha x + \beta m = (\alpha x_1 + \beta m_1, \cdots \alpha x_n + \beta m_n), (m, x) = m_1 x_1 + \cdots + m_n x_n \text{ and } |x| = (x, x)^{\frac{1}{2}}.$$

As in [1], a multiple trigonometric series $\sum a_m e^{i(m,x)}$, where m represents a lattice point and a_m is an arbitrary complex number, will be said to be Abel summable at the point x to the value $L(x)$ if

$$\sum a_m \exp[i(m, x) - |m| z] \to L(x) \text{ as } z \to 0.$$

The series $\sum a_m e^{i(m,x)}$ will be called a real-valued series if $\bar{a}_m = a_{-m}$, and a series of class (U') if $\sum' a_m |m|^{-2} e^{i(m,x)}$, where $m \neq 0$, is the Fourier series of a continuous periodic function.

The open sphere of radius t and center x will be denoted in this paper by $D_n(x, t)$; the surface of this sphere by $C_n(x, t)$. The torus or fundamental semi-closed cube $\{x; -\pi < x_i \leqq \pi, i = 1, \cdots, n\}$ will be designated

* Received December 1, 1955.

[1] National Science Foundation Fellow.

by Ω_n. $x+\Omega_n$ will be the set $\{p; p-x \varepsilon \Omega_n\}$. Z contained in Ω_n will be said to be a closed set in Ω_n if it contains all its limit points in E_n except those which lie on the faces $x_i=-\pi$, $i=1,\cdots,n$. Also these latter limit points when considered as points on the faces $x_i=\pi$ do lie in Z. In other words Z will be said to be closed in Ω_n if it is closed in a torus sense. It is clear that if Z is such a set and also of capacity zero its closure in E_n is also of capacity zero.

Given $F(x)$ a real-valued function in $D_n(x_0, t_0)$ which is integrable on $C_n(x_0, t)$ for $0<t\leqq t_0$, we shall designate the mean value of F on this latter surface by $L(F, x_0, t)$. Thus

$$L(F, x_0, t) = \omega_n^{-1} \int_{C_n(0,1)} F(x_0 + tx)\, dS_{n-1}(x)$$

where $\omega_n = 2\pi^{\frac{1}{2}n}/\Gamma(\frac{1}{2}n)$ is the $(n-1)$-dimensional volume of $C_n(0,1)$ and $dS_{n-1}(x)$ is the $(n-1)$-dimensional volume element of $C_n(0,1)$.

We say that $F(x)$ has an r-th generalized Laplacian at the point x_0 equal to α_r if

$$L(F, x_0, t) = \Gamma(n/2) \sum_{j=0}^{r} t^{2j}\alpha_j / 2^{2j} j! \Gamma(j + n/2) + o(t^{2r})$$

where the α_j $(j=0,\cdots,r)$ are constants. This r-th generalized Laplacian will be designated by $\Delta_r F(x_0)$, and it is clear that if $\Delta_r F(x_0)$ exists, then $\Delta_s F(x_0)$ exists for $0\leqq s\leqq r$. By [4, p. 261] if $F(x)$ is in class C^{2r} in $D_n(x_0, t)$ for some $t>0$, then $\Delta_r F(x_0)$ exists and equals $\Delta^r F(x_0)$, where Δ^r stands for the usual Laplace operator iterated r-times.

We set $\nabla(F, x_0, t) = L(F, x_0, t) - F(x_0)$ and

$$\psi^* F(x_0) = \limsup_{t\to 0} 2n \nabla(F, x_0, t)/t^2 \tag{1}$$

the upper first generalized Laplacian of F at the point x_0. Replacing lim sup by lim inf in (1) we have a similar definition for the lower first generalized Laplacian $\psi_* F(x_0)$. If $\psi^* F(x_0) = \psi_* F(x_0)$ is finite, then F has a first generalized Laplacian at the point x_0.

Throughout this paper μ will always designate the value $(n-2)/2$, and $J_\mu(t)$ will stand for the Bessel function of the first kind of order μ.

The capacity of a set in E_n will refer to the logarithmic capacity if $n=2$ and n-dimensional Newtonian capacity if $n\geqq 3$. In this connection we construct the function $\Phi(x)$ defined in Ω_n as

$$\Phi(x) = 2\pi \log |x|^{-1} \text{ for } n=2 \text{ and } \Phi(x) = (2\pi)^n [\omega_n (n-2)]^{-1} |x|^{-(n-2)}$$
$$\text{for } n\geqq 3.$$

$\Phi(x)$ is then defined in the rest of E_n by periodicity. In other words, if $\eta = 2\pi(j_1, \cdots, j_n)$, where j_k $(k = 1, \cdots, n)$ are integers, then $\Phi(x+\eta) = \Phi(x)$.

The function whose Fourier series is $\sum' e^{i(m,x)} | m |^{-2}$, where $m \neq 0$, will be designated in this paper by $G(x)$. The properties of $G(x)$, which are the key to the whole paper, will be discussed in Section 4.

If $f(x)$ is integrable in a bounded domain R', by $\Phi * f$ we shall mean $\int_{R'} \Phi(x-y) f(y)\, dy$. By $F = \Phi * f$ will be meant that

$$\int_{R'} | F(x) - \Phi * f(x) |\, dx = 0.$$

$H_r(x)$ will be said to be harmonic of order r in R' if $H_r(x)$ is in class C^∞ and $\Delta^r H_r(x) = 0$.

If $f(x)$ is integrable on Ω_n, then $S[f]$ will designate the Fourier series of f.

$\bar{R}$ will designate the closure of R in E_n.

3. Statement of main results. We shall prove the following theorems:

THEOREM 1. *Let $f(x)$ and $\Delta_j f(x)$ $(j = 1, \cdots, r-1)$ be continuous in a bounded domain R contained in E_n and let Z be a closed and bounded set of capacity zero contained in E_n. Suppose that $\Delta_r f(x)$ exists in $R - RZ$ and is integrable on R. Then in every subdomain R' whose closure is contained in R,*

$$f = \underbrace{P * \cdots * P}_{r} * \Delta_r f + H_r,$$

where $H_r(x)$ is harmonic of order r in R' and $P(x) = -(2\pi)^{-1} \log | x |^{-1}$ if $n = 2$ and $= -[(n-2)\omega_n]^{-1} | x |^{-(n-2)}$ if $n \geqq 3$.

Remark. The condition that f and $\Delta_j f$ be continuous in the above theorem is necessary as can be seen from the following example in E_2. Set $f(x) = x_1^2$ if $x_1 \geqq 0$ and $= -x_1^2$ if $x_1 \leqq 0$. Then $\Delta_2 f(x) = 0$ for all x in E_2. But $f(x)$ is clearly not harmonic of order 2.

THEOREM 2. *Let $F(x)$ and $g(x)$ be two real-valued periodic functions of period 2π in each variable, with $F(x)$ continuous in E_n and $g(x)$ integrable on the torus Ω_n. Also let Z be a closed (in the torus sense) set of capacity zero contained in Ω_n. Suppose that*

(i) $\psi^* F(x) > -\infty$, $\psi_* F(x) < +\infty$ *for x in $\Omega_n - Z$;*

(ii) $g(x) \leqq \psi^* F(x)$ *in* Ω_n.

Then

a) $\Delta_1 F(x)$ *exists at almost all points* x *of* Ω_n *and is in* L_1 *on* Ω_n;

b) *at all points* x *for which* $\int_{\Omega_n} |\Delta_1 F(y)|\,|\Phi(x-y)|\,dy < \infty$, *we have*

$$F(x) = -(2\pi)^{-n} \int_{\Omega_n} \Delta_1 F(y)\, G(x-y)\, dy + (2\pi)^{-n} \int_{\Omega_n} F(y)\, dy.$$

THEOREM 3. *Given the multiple trigonometric series* $\sum a_m e^{i(m,x)}$ *where the* a_m *are arbitrary complex numbers. Let* Z *be a closed (in the torus sense) set of capacity zero contained in* Ω_n. *Suppose that*

(i) *the series is of class* (U');

(ii) *the series is Abel summable to* $f(x)$ *almost everywhere where* $f(x)$ *is in* L_1 *on* Ω_n;

(iii) $\limsup_{z \to 0} |\sum a_m e^{i(m,x)-|m|z}| < \infty$ *in* $\Omega_n - Z$.

Then the series is the Fourier series of $f(x)$.

4. Green's function on the torus. Before proving Theorem 1, it is necessary to investigate the properties of the function $G(x)$ whose Fourier series is $\sum' e^{i(m,x)} |m|^{-2}$. We first obtain the following lemma with $\Phi(x)$ as in Section 2.

LEMMA 1. $G(x) = \Phi(x) + H^*(x)$ *where* $H^*(x)$ *is continuous in* E_n.

Let us set λ_m^{-1} equal to the m-th Fourier coefficient of $\Phi(x)$. Then we prove the lemma by showing that $\sum_{m \neq 0} |\lambda_m^{-1} - |m|^{-2}| < \infty$.

For $n = 2$, this already has been shown in [7]. For $n \geqq 3$, we observe by Green's second identity that for $\mu = (n-2)/2$ and $m \neq 0$ that

$$(2) \qquad [\omega_n(n-2)]^{-1} \int_{\Omega_n - D_n(0,\epsilon)} |x|^{-(n-2)} e^{i(m,x)} dx$$
$$= 2^\mu \Gamma(\mu+1) J_\mu(|m|\epsilon)(|m|\epsilon)^{-\mu} |m|^{-2} + o(1)$$
$$+ K|m|^{-2} \sum_{j=1}^{n} \int_{\Omega_{n-1}} \frac{\exp[i(m_1 x_1 + \cdots + m_j x_j^* + \cdots + m_n x_n)]}{(x_1^2 + \cdots + \pi^2 + \cdots + x_n^2)^{n/2}}$$
$$\times \cos m_j \pi\, dx_1 \cdots dx_j^* \cdots dx_n$$

where K is a constant independent of m and ϵ and $m_j x_j^*$ stands for the deletion of $m_j x_j$.

Integration by parts $(n-1)$ times shows us that

$$(3)\quad \Big|\int_{\Omega_{n-1}} \exp[i(m_1x_1+\cdots+m_jx_j^*+\cdots+m_nx_n)] \times(x_1^2+\cdots+\pi^2+\cdots+x_n^2)^{-n/2}dx_1\cdots dx_j^*dx_n\Big| \leqq K_1[(|m_1|+1)\cdots(|m_j|+1)^*\cdots(|m_n|+1)]^{-1}$$

where K_1 is another constant independent of m.

We therefore conclude from (2) and (3) that for $m\neq 0$

$$(3')\quad |\lambda_m^{-1}-|m|^{-2}| \leqq K_2|m|^{-2}\sum_{j=1}^{n}[(|m_1|+1)\cdots \times(|m_j|+1)^*\cdots(|m_n|+1)]^{-1}$$

where K_2 is another constant independent of m.

We next observe that for $m\neq 0$ there is a K_3 independent of m such that

$$[(|m_1|+1)\cdots(|m_n|+1)]^{(n+1)/n} \leqq K_3|m|^2(|m_1|+1)\cdots(|m_j|+1)^*\cdots(|m_n|+1).$$

The lemma then follows from (3′) and the fact that

$$\sum_m[(|m_1|+1)\cdots(|m_n|+1)]^{-(n+1)/n}<\infty.$$

LEMMA 2. *$H^*(x)$, defined in Lemma 1, is in $C^{(\infty)}$ in the interior of Ω_n. and furthermore $\Delta H^*(x)=1$ in this domain.*

By [2, Theorem 6], $\Delta G(x)=1$ in the interior of Ω_n provided $x\neq 0$. Since for such x, $\Delta\Phi(x)=0$, we have that $\Delta[H^*(x)-|x|^2/2n]=0$ for x in the interior of Ω_n and $x\neq 0$. But $H^*(x)-|x|^2/2n$ is continuous at the origin by Lemma 1. Consequently $H^*(x)-|x|^2/2n$ is also harmonic in a neighborhood of the origin, and the lemma is proved.

LEMMA 3. *Let $f(x)$ be a continuous function of period 2π in each variable and such that $\int_{\Omega_n} f(x)dx=0$. Suppose that $S[f]=\sum' a_me^{i(m,x)}$ and $S[F]=\sum'(-1)^re^{i(m,x)}a_m/|m|^{2r}$, where $F(x)$ is taken to be continuous. Then $\Delta_rF(x)=f(x)$ for every x in Ω_n.*

We shall give the proof for $n\geqq 3$; a similar proof holding for $n=2$. From periodicity there is no loss in generality in just proving the

theorem for the origin. We first prove the theorem, therefore, at this point with $r=1$. Then

$$F(x) = -(2\pi)^{-n}\int_{\Omega_n}[G(x-y)-|x|^2/2n]f(y)dy$$

and with t sufficiently small,

$$\nabla(F,0,t) = -(2\pi)^{-n}\int_{\Omega_n}\nabla[G(x-y)-|x|^2/2n,0,t]f(y)dy.$$

But for t small and fixed y in $\Omega_n - \bar{D}_n(0,t)$, the function $G(x-y)-|x|^2/2n$ is harmonic in $D_n(0,t+\epsilon)$, where ϵ is a small positive number depending on y. Therefore,

$$\nabla(F,0,t) = -(2\pi)^{-n}\int_{D_n(0,t)}\nabla[G(x-y)-|x|^2/2n,0,t]f(y)dy.$$

However $G(x-y)-|x|^2/2n = \Phi(x-y)+[H(x-y)-|x|^2/2n]$, where, by Lemma 2, the expression in brackets is harmonic as a function of x for fixed y when t is small and x and y are in $D_n(0,t)$. We conclude that

$$\begin{aligned}(4)\qquad \nabla(F,0,t) &= -(2\pi)^{-n}\int_{D_n(0,t)}\nabla[\Phi(x-y),0,t]f(y)dy \\ &= [\omega_n(n-2)]^{-1}\int_{D_n(0,t)}[|y|^{2-n}-t^{2-n}]f(y)dy.\end{aligned}$$

The proof of the lemma for the case $r=1$ follows immediately from (4) and the definition of the first generalized Laplacian.

Let us assume now that the theorem is true for $1 \leqq s \leqq r-1$. Then letting g be the continuous function whose Fourier series is

$$\Sigma'(-1)^{r-1}a_m e^{i(m,x)}/|m|^{2(r-1)},$$

we have by assumption that

$$(5)\quad L(g,0,t) = \alpha_0 + \alpha_1 t^2 + \cdots + \alpha_{r-2}t^{2(r-2)} + f(0)t^{2(r-1)}/\beta_{r-1} + o(t^{2(r-1)})$$

where $\beta_\nu = 2^{2\nu}\nu!\Gamma(\nu+n/2)/\Gamma(n/2)$ and the α_j are constants.

From (4) we obtain that

$$(6)\qquad \nabla(F,0,t) = (n-2)^{-1}\int_0^t L(g,0,\rho)[\rho^{2-n}-t^{2-n}]\rho^{n-1}\,d\rho.$$

Putting (5) in (6) and integrating, we obtain

$$(7)\qquad \nabla(F,0,t) = \sum_{j=1}^{r-1}\gamma_j t^{2j} + f(0)t^{2r}/\beta_{r-1}2r(n+2r-2) + o(t^{2r}).$$

But $\beta_\nu 2(\nu+1)(n+2\nu) = \beta_{\nu+1}$, and the lemma is proved by (7) and the very definition of the r-th generalized Laplacian.

5. Generalized Laplacians and the Abel summation of trigonometric series. In this section we prove a lemma which coupled with Theorem 2 will enable us to prove Theorem 3. This lemma will also be very useful in the proof of Theorem 1. (For the 1-dimensional analogue of this lemma see [9, p. 398].

LEMMA 4. *Let $f^*(x)$ and $f_*(x)$ be the upper and lower Abel sums of the real-valued multiple trigonometric series $S = \sum_{m\neq 0} a_m e^{i(m,x)}$. Suppose that $C - \sum_{m\neq 0} a_m |m|^{-2} e^{i(m,x)}$ is the Fourier series of a continuous periodic function $F(x)$. Then for all x we have*

a) $\psi_* F(x) \leqq f^*(x)$; b) $f_*(x) \leqq \psi^* F(x)$.

We only need prove a) for b) will then follow from considering $-F(x)$. Also from the assumptions of the theorem, we see that it is sufficient to prove a) only for $x = 0$. Furthermore if $F(0) \neq 0$ we can make it so by adding a constant since $\psi^* c = \psi_* c = 0$.

To prove a), it is sufficient to show that if $\psi_* F(0) > p$ then $f^*(0) \geqq p$. It is clear that if $p = -\infty$, a) is proved. If $p \neq -\infty$, we can assume that it is non-negative for otherwise we can consider $F(x) - p[1 - \cos(m_1, x)/|m_1|^2]$ where m_1 is a fixed integral lattice point different from the origin.

Let us assume then that $f^*(0) < 0$ and show this contradicts the fact that $\psi_* F(0) = h > 0$. Setting $F(z,x) = C - \sum_{m\neq 0} a_m e^{i(m,x)-|m|z} |m|^{-2}$ and $f(z,x) = \sum_{m\neq 0} a_m e^{i(m,x)-|m|z}$, we see that for $z > 0$, $F(z,x)$ is an harmonic function in the variables (z,x) and that $F_{zz}(z,x) + f(z,x) = \Delta F(z,x) = 0$. Consequently there is an $\epsilon > 0$ such that

$$F_{zz}(z,0) > 0 \text{ for } 0 < z \leqq \epsilon. \tag{8}$$

Furthermore since the continuity of $F(x)$ implies that $\lim_{z\to 0} F(z,0) = 0$, we have by the mean-value theorem that for every z in $0 < z \leqq \epsilon$ there is an s such that $0 < s < z$ and such that $F(z,0)/z = F'(s,0)$ where $F'(z,0) = F_z(z,0)$. (8) then implies that $F(z,0)/z - F(s,0)/s = F'(s,0) - F'(t,0) > 0$ where $0 < t < s < z$. Therefore to obtain a contradiction to the fact that $f^*(0) < 0$, it is sufficient to show $\limsup_{z\to 0} \partial[F(z,0)/z]/\partial z < 0$ or what is the same thing that

$$\liminf_{z\to 0} \partial[-F(z,0)/z]/\partial z > 0. \tag{9}$$

We shall now show that (9) holds. For by [1, p. 176]

$$F(z,0)/z = K_n \int_0^\infty L(F,0,t)\,(z^2+t^2)^{-(n+1)/2} t^{n-1}\,dt$$

where K_n is a positive constant.

But since $F(x)$ is a continuous periodic function, it is not difficult to see that for every $z>0$, there is an open interval containing z such that

$$\lim_{R\to\infty} \int_0^R zL(F,0,t)\,(z^2+t^2)^{-(n+3)/2} t^{n-1}\,dt$$

is uniformly convergent in this interval. Therefore

$$(10)\quad \partial[-F(z,0)/z]/\partial z = (n+1)K_n \int_0^\infty zL(F,0,t)\,(z^2+t^2)^{-(n+3)/2} t^{n-1}\,dt.$$

From the fact that $\psi_* F(0) = h > 0$, we have that there is a $\delta > 0$ such that for $0 < t \leqq \delta$, $L(F,0,t) > ht^2/4n$. The fact in conjunction with the observation that $\lim_{z\to 0} z \int_\delta^\infty t^{n-1}[z^2+t^2]^{-(n+3)/2}\,dt = 0$, tells us from (10) that

$$(11)\quad \liminf_{z\to 0} \partial[-F(z,0)/z]/\partial z > hK_n' \liminf_{z\to 0} \int_0^\delta z(z^2+t^2)^{-(n+3)/2} t^{n+1}\,dt,$$

where K_n' is a positive constant. But

$$\lim_{z\to 0} \int_0^\delta z(z^2+t^2)^{-(n+3)/2} t^{n+1}\,dt = \int_0^\infty t^{n+1}(1+t^2)^{-(n+3)/2}\,dt > 0,$$

and we have shown that (9) holds.

By $\Delta^r S[f]$ we shall mean the trigonometric series that one obtains by formally applying the Laplace operator r-times to $S[f]$.

We now state a lemma concerning such series whose proof follows immediately from [8, p. 225] for the two-dimensional case and from a similar theorem for the n-dimensional case.

LEMMA 5. *If $\Delta_r f(x_0)$ exists, $\Delta^r S[f]$ is at the point x_0 Abel summable to $\Delta_r f(x_0)$.*

6. Proof of Theorem 1. Since R is assumed to be a bounded domain, we shall assume further that its closure is contained in the interior of Ω_n. From the proof it will be apparent that this additional assumption will cause no loss of generality. Also we shall assume from the start that $f(x)$ is real-valued with no loss of generality.

Since $\bar{R}' \subset R$, we can insert three other domains between them with the same property as R', i.e. $\bar{R}' \subset R'' \subset \bar{R}'' \subset R''' \subset \bar{R}''' \subset R^{\text{iv}} \subset \bar{R}^{\text{iv}} \subset R$ and we can form the localizing function $\lambda(x)$ for R''' and R^{iv}. Then $\lambda(x) = 1$ for x in R''' and $= 0$ for x in Ω_n and not in R^{iv}, and furthermore $\lambda(x)$ is in class C^∞ for x in Ω_n.

By [5, Theorem 1], the theorem is true for $r = 1$. Let us assume then that the theorem is true for $1 \leqq s \leqq r-1$ and let us set $g(x) = \lambda(x)\Delta_{r-1}f(x)$, $S[g] = \sum_m b_m e^{i(m,x)}$, and $S[F] = \sum_{m \neq 0} (-1)^{r-1} b_m e^{i(m,x)}/|\, m \,|^{2(r-1)}$ where $F(x)$ is taken to be continuous. Then by Lemma 3, $\Delta_{r-1}(F - b_0 \,|\, x \,|^{2(r-1)}/k_{r-1}) = g(x)$ where $\Delta^j \,|\, x \,|^{2j} = k_j$, and consequently in R''', $\Delta_{r-1}(F - f - b_0 \,|\, x \,|^{2(r-1)}/k_{r-1} = 0$. Therefore by the inductive assumption

$$F(x) = f(x) + H_r(x) \text{ for } x \text{ in } R'' \tag{12}$$

where $H_r(x)$ is harmonic of order r in R''.

From (12) we see that in $R'' - R''Z$, $\Delta_r F(x) = \Delta_r f(x)$. Consequently by Lemma 5, $-\sum_m b_m \,|\, m \,|^2 e^{i(m,x)}$ is Abel summable to $\Delta_r f(x)$ in $R'' - R''Z$. But then by Lemma 4, $\psi^* g \geqq \Delta_r f \geqq \psi_* g$ in $R'' - R''Z$. Since g is a continuous function in R'', the conditions of [5, Theorem 1] are satisfied. So in R'

$$g = P * \Delta_r f + H_1(x) \tag{13}$$

where $H_1(x)$ is harmonic in R'.

Since $\Delta_{r-1}(F - b_0 \,|\, x \,|^{2(r-1)}/k_{r-1}) = g$, we have by the inductive assumptions that in R'

$$F = \underbrace{P * \cdots * P}_{r-1} * g + H_{r-1} + b_0 \,|\, x \,|^{2(r-1)}/k_{r-1}. \tag{14}$$

From (12), (13), and (14), we obtain that, in R',

$$f = \underbrace{P * \cdots * P}_{r} * \Delta_r f + H_{R'}$$

where $H_{R'}(x)$ is harmonic of order r, and the proof to the theorem is complete.

7. Proof of Theorem 2. By [5, Theorem 1 and 1.3.2] for almost all x in any bounded domain R contained in the plane

$$F(x) = -(2\pi)^{-1} \int_R \Delta_1 F(y) \log |\, x - y \,|^{-1} \, dy + h_1(x) \tag{15}$$

where $h_1(x)$ is harmonic in R. In a similar manner it can be shown that for any bounded domain R contained in E_n $(n \geqq 3)$ and for almost all x in R

$$F(x) = -[\omega_n(n-2)]^{-1} \int_R \Delta_1 F(y) |x-y|^{-(n-2)} dy + h_1(x) \tag{16}$$

where $h_1(x)$ is harmonic in R. In both cases $\Delta_1 F(y)$ is in L_1 on R.

We conclude, therefore, that for any x_0 in E_n there is a continuous function $F_{x_0}(x)$ which for almost all x in $D_n(x_0, \pi/4)$ is such that

$$F_{x_0}(x) = -(2\pi)^{-n} \int_{D_n(x_0,\pi/4)} \Phi(x-y)\Delta_1 F(y)\, dy \tag{17}$$

and. furthermore, for all x in $D_n(x_0, \pi/4)$

$$\Delta[F(x) - F_{x_0}(x)] = 0. \tag{18}$$

We also observe that $\int_{\Omega_n + x_0 - D_n(x_0,\pi/4)} \Phi(x-y)\Delta_1 F(y)\, dy$ is continuous in $D_n(x_0, \pi/4)$ and that for almost all x in E_n,

$$\int_{\Omega_n + x_0} \Phi(x-y)\Delta_1 F(y)\, dy = \int_{\Omega_n} \Phi(x-y)\Delta_1 F(y)\, dy.$$

These facts in conjunction with (17) and (18) tell us that there exists a continuous function $F_1(x)$ which is also periodic of period 2π in each variable such that for almost all x in E_n,

$$F_1(x) = -(2\pi)^{-n} \int_{\Omega_n} G(x-y)\Delta_1 F(y)\, dy. \tag{19}$$

But then from (18), Lemma 2, and the fact that $\Delta G(x) = 1$ for $x \neq 0 \bmod \Omega_n$, we obtain that, for all x in $D_n(x_0, \pi/4)$,

$$\Delta[F - F_1] = (2\pi)^{-n} \int_{\Omega_n} \Delta_1 F(y)\, dy.$$

Since x_0 is arbitrary and $F - F_1$ is a periodic continuous function, it must be that $F(x) = F_1(x) + \text{constant}$ on Ω_n. That the constant agrees with that of the theorem follows from the fact that the integral of $G(x)$ over Ω_n is zero.

Part b) of the theorem follows from the fact that by [5, Theorem 1] equality holds in (17) at all points at which b) is satisfied. But then equality also holds at those points in (19), and the theorem is proved.

8. Proof of Theorem 3. There is no loss of generality if from the start we assume that the given trigonometric series is a real-valued series. For suppose the theorem is proved under this assumption. Then it is clear that both of the series $\sum(a_m + \bar{a}_{-m})e^{i(m,x)}$ and $\sum i(a_m - \bar{a}_{-m})e^{i(m,x)}$ are of class (U'), are Abel summable to $f(x) + \bar{f}(x)$ and $i[f(x) - \bar{f}(x)]$ almost everywhere respectively, are real-valued trigonometric series, and satisfy (iii). We would then have

$$2a_m = (a_m + \bar{a}_{-m}) + (a_m - \bar{a}_{-m})$$

$$= (2\pi)^{-n} \int_{\Omega_n} e^{i(m,x)}[f(x) + \bar{f}(x) + f(x) - \bar{f}(x)]dx$$

and the theorem would be proved in the more general case.

We also see there is no loss in generality in asuming that $a_0 = 0$. For if the theorem is proved with constant term equal to zero, we would then have $(2\pi)^{-n} \int_{\Omega_n} [f(x) - a_0]dx = 0$, and consequently the theorem is true in the more general case.

With the given series a real-valued series, with $a_0 = 0$, and with $F(x)$ the continuous periodic function whose Fourier series is $-\sum_{m \neq 0} a_m e^{i(m,x)} | m |^{-2}$, we then have by Lemma 4 and assumption (ii) of this theorem that $\psi^* F(x) \geqq f(x)$ almost everywhere. Setting $g(x) = \min[\psi^* F(x), f(x)]$, we have that $g(x)$ is integrable on Ω_n. But then by Theorem 2, $\Delta_1 F(x)$ exists almost everywhere in Ω_n, and furthermore by Lemma 5 and (ii) of the present theorem, $\Delta_1 F(x) = f(x)$ almost everywhere. We conclude, consequently from Theorem 2, that for almost all x in Ω_n,

$$F(x) = -(2\pi)^{-n} \int_{\Omega_n} f(y) G(x - y)dy. \tag{20}$$

Observing that $\int_{\Omega_n} |f(y)|\, dy \int_{\Omega_n} |G(x - y)|\, dx < \infty$, we have that for $m \neq 0$, by Fubini's theorem and (20),

$$(2\pi)^n a_m | m |^{-2} = \int_{\Omega_n} e^{-i(m,x)} F(x)dx$$

$$= \int_{\Omega_n} e^{-i(m,y)} f(y)dy[(2\pi)^{-n} \int_{\Omega_n} e^{-i(m,x-y)} G(x - y)dx]$$

$$= \int_{\Omega_n} e^{-i(m,y)} f(y) | m |^{-2} dy.$$

This last fact in conjunction with the observation that by [1] the Fourier series of $f(x)$ is Abel summable to $f(x)$ almost everywhere tells us that $\int_{\Omega_n} f(x)dx = 0$, which concludes the proof to the theorem.

INSTITUTE FOR ADVANCED STUDY AND RUTGERS UNIVERSITY.

REFERENCES.

[1] S. Bochner, "Summation of multiple Fourier series by spherical means," *Transactions of the American Mathematical Society*, vol. 40 (1936), pp. 175-207.

[2] ———, "Zeta functions and Green's functions for linear partial differential operators of elliptic type with constant coefficients," *Annals of Mathematics*, vol. 57 (1953), pp. 32-56.

[3] M. T. Cheng, "Uniqueness of multiple trigonometric series," *Annals of Mathematics*, vol. 52 (1950), pp. 403-416.

[4] R. Courant and D. Hilbert, *Methoden der mathematischen Physik II*, Berlin, 1937.

[5] W. Rudin, "Integral representation of continuous functions," *Transactions of the American Mathematical Society*, vol. 68 (1950), pp. 278-286.

[6] V. L. Shapiro, "An extension of results in the uniqueness theory of double trigonometric series," *Duke Mathematical Journal*, vol. 20 (1953), pp. 359-365.

[7] ———, "Logarithmic capacity of sets and double trigonometric series," *Canadian Journal of Mathematics*, vol. 6 (1954), pp. 582-592.

[8] ———, "Circular summability C of double trigonometric series," *Transactions of the American Mathematical Society*, vol. 76 (1954), pp. 223-233.

[9] A. Zygmund, *Trigonometric series*, Warsaw, 1935.

THE FIELD OF DEFINITION OF A VARIETY.*

By ANDRÉ WEIL.

Let V be a variety, defined over an overfield K of a groundfield k. Consider the following problems:

(P) Among the varieties, birationally equivalent to V over K, find one which is defined over k.

(P′) Among the varieties, birationally and biregularly equivalent to V over K, find one which is defined over k.

Problems of these types arise, for instance, in Chow's recent work on abelian varieties over function-fields ([1]), in my work on algebraic groups ([4]), and also in the unpublished work of Shimura and of Taniyama on complex multiplication. Criteria for those problems to have a solution are implicitly contained in Chow's paper ([1]) and in Lang's subsequent note on a related subject ([2]); the purpose of the present paper is to develop them more explicitly.

Without restricting the generality of the problem, we may assume that K is finitely generated over k; we shall make the restrictive assumption that it is separable over k. Then it is a regular extension of the algebraic closure k' of k in K; and k' is a separably algebraic extension of k of finite degree. Thus it will be enough for our purpose to discuss (P) and (P′), firstly when K is separably algebraic over k, and secondly when it is regular over k.

As usual, we do not distinguish between mappings and their graphs. In particular, we do not distinguish between a birational correspondence T between two varieties V and W (this being defined as a subvariety of $V \times W$ which satisfies certain conditions) and the mapping of V into W determined by T. The inverse mapping, T^{-1}, is then a mapping of W into V or also a birational correspondence between W and V. To prevent misunderstandings, I take this opportunity for pointing out that (by abuse of language) I call T *everywhere biregular* only when T is biregular at every point of V and T^{-1} is so at every point of W; when that is so, T might more suitably be

---

* Received January 16, 1956.

called an *isomorphism* between V and W, and a *k-isomorphism* if k is a field of definition for V, W and T.

Section I. Separably Algebraic Extensions of the Groundfield.

1. Let k be a separably algebraic extension of a groundfield k_0, of finite degree n. Call $\mathfrak{J}$ the set of all distinct isomorphisms of k (over k_0, i.e., leaving all elements of k_0 invariant) into the algebraic closure $\bar{k}_0$ of k_0; $\mathfrak{J}$ consists of n distinct isomorphisms, including the identity automorphism ϵ of k. If $\sigma \in \mathfrak{J}$, and if ω is any isomorphism over k_0 of an overfield of k^σ, we denote by $\sigma\omega$ the isomorphism of k defined by putting $\xi^{\sigma\omega} = (\xi^\sigma)^\omega$ for every $\xi \in k$.

Let V be a variety, defined over k; assume that there is a variety V_0, defined over k_0, and a birational correspondence f, defined over k, between V_0 and V. Then, for $\sigma \in \mathfrak{J}$, $\tau \in \mathfrak{J}$, the mapping $f_{\tau,\sigma} = f^\tau \circ (f^\sigma)^{-1}$ is a birational correspondence between V^σ and V^τ. We now modify our problem (P) as follows:

(A) *Let k be a separably algebraic extension of k_0 of finite degree; let $\mathfrak{J}$ be the set of all distinct isomorphisms of k into $\bar{k}_0$. Let V be a variety, defined over k; for each pair (σ, τ) of elements of $\mathfrak{J}$, let $f_{\tau,\sigma}$ be a birational correspondence between V^σ and V^τ. Find a variety V_0, defined over k_0, and a birational correspondence f, defined over k, between V_0 and V, such that $f_{\tau,\sigma} = f^\tau \circ (f^\sigma)^{-1}$ for all $\sigma \in \mathfrak{J}$, $\tau \in \mathfrak{J}$.*

THEOREM 1. *Problem* (A) *has a solution if and only if the $f_{\tau,\sigma}$ are defined over a separably algebraic extension of k_0 and satisfy the following conditions:*

(i) $f_{\tau,\rho} = f_{\tau,\sigma} \circ f_{\sigma,\rho}$ *for all ρ, σ, τ in $\mathfrak{J}$;*

(ii) $f_{\tau\omega,\sigma\omega} = (f_{\tau,\sigma})^\omega$ *for all σ, τ in $\mathfrak{J}$ and all automorphisms ω of $\bar{k}_0$ over k_0.*

Moreover, when that is so, the solution is unique, up to a birational transformation on V_0, defined over k_0.

The conditions are obviously necessary. If the problem has two solutions (V_0, f) and (V_0', f'), put $F = f'^{-1} \circ f$; this is a birational correspondence between V_0 and V_0', defined over k. Writing that (V_0, f) and (V_0', f') are solutions of (A), we find $F^\sigma = F^\tau$ for all σ, τ; thus, F is invariant under all automorphisms of $\bar{k}_0$ over k_0; therefore it is defined over k_0; this proves the unicity assertion in Theorem 1.

Now assume that (i), (ii) are fulfilled; then, if σ, τ are in $\mathfrak{A}$ and ω is an automorphism of $\bar{k}_0$ over the compositum of k^σ and k^τ, (ii) shows that $f_{\tau,\sigma}$ is invariant under ω; if $f_{\tau,\sigma}$ is defined over a separably algebraic extension of k_0, this implies that it is defined over the compositum of k^σ and k^τ. Let x be a generic point of V over k; for each $\sigma \in \mathfrak{A}$, put $x_\sigma = f_{\sigma,\epsilon}(x)$, ϵ being the identity automorphism of k. If we put $\rho = \sigma$ in (i), we see that $f_{\sigma,\sigma}$ is the identity mapping of V^σ; therefore we have $x_\epsilon = x$.

Let K be the compositum of the fields k^σ, for all $\sigma \in \mathfrak{A}$; this is a Galois extension of k_0; call Γ its Galois group. Take any $\omega \in \Gamma$; as x and $x_{\epsilon\omega}$ are generic points of V and V^ω, respectively, over K, there is one and only one isomorphism ω^* of $K(x)$ onto $K(x_{\epsilon\omega})$ which induces ω on K and maps x onto $x_{\epsilon\omega}$. As $f_{\epsilon\omega,\epsilon}$ is a birational correspondence, defined over K, we have $K(x) = K(x_{\epsilon\omega})$, so that ω^* is an automorphism of $K(x)$. Applying (ii) to any extension of ω to an automorphism of $\bar{k}_0$, we get:

$$(x_\sigma)^{\omega^*} = f_{\sigma\omega,\epsilon\omega}(x_{\epsilon\omega}) ;$$

putting $\rho = \epsilon$ and $\sigma = \epsilon\omega$ in (i), we find that the right-hand side of this relation is $x_{\sigma\omega}$; therefore ω^* maps x_σ onto $x_{\sigma\omega}$ for all ω. From this it immediately follows that the mapping $\omega \to \omega^*$ is a homomorphism of Γ into the group of all automorphisms of $K(x)$, and more precisely an isomorphism of Γ onto a group Γ^* of automorphisms of $K(x)$. Call $k_0(y)$ the field consisting of those elements of $K(x)$ which are invariant under Γ^*; it is finitely generated over k_0 ([**4**], App., Prop. 3). As $K(x)$ is regular, hence separable, over K, and K is separable over k_0, $K(x)$ is separable over k_0 (Bourbaki, *Alg.*, Chap. V, § 7, no. 4, Prop. 7); hence $k_0(y)$ is separable over k_0. Any element of the algebraic closure of k_0 in $k_0(y)$ must be in the algebraic closure of K in $K(x)$, which is K since $K(x)$ is regular over K; as such an element is invariant under Γ^*, it must then be invariant under Γ, and so it must be in k_0. Thus we have proved that $k_0(y)$ is regular over k_0. Call V_0 the locus of y over k_0.

If an element ω of Γ induces the identity on k, ω^* leaves x invariant; as Γ^* is the Galois group of $K(x)$ over $k_0(y)$, this implies that $k(x) \subset k(y)$, so that we may write $x = f(y)$, where f is a mapping of V_0 into V, defined over k. We have $K(x) \subset K(y)$, hence $K(x) = K(y)$ since $K(y)$ is contained in $K(x)$ by definition. This shows that f is a birational correspondence between V_0 and V. Transforming the relation $x = f(y)$ by any ω^* in Γ^*, and calling σ the isomorphism of k induced on k by ω^*, we get $x_\sigma = f^\sigma(y)$, hence $f_{\sigma,\epsilon} = f^\sigma \circ f^{-1}$; by (i), this shows that (V_0, f) is a solution of our problem.

2. In the introduction, we formulated, in addition to the problem (P), a more precise problem (P′). We may modify the problem (A) similarly, by requiring f to be everywhere biregular; call (A′) this modified problem. For (A′) to have a solution, it is obviously necessary that the $f_{\tau,\sigma}$ should be everywhere biregular and satisfy the conditions in Theorem 1.

Assume that this is so; let (V_0, f) be a solution of (A). Then, if (V_0', f') is a solution of (A′), the unicity of the solution of (A) shows that we must have $f' = f \circ F^{-1}$, where F is a birational correspondence between V_0 and V_0', defined over k_0. Thus problem (A′) may be reformulated as follows:

(B) *Let k and k_0 be as in problem* (A); *let V and V_0 be varieties, respectively defined over k and over k_0; let f be a birational correspondence, defined over k, between V_0 and V. Find a variety V_0' and a birational correspondence F between V_0 and V_0', both defined over k_0, such that the birational correspondence $f \circ F^{-1}$ between V_0' and V is everywhere biregular.*

It is obvious that, if (B) has a solution, this is unique up to a k_0-isomorphism; therefore the same is true for (A′). If (B) has a solution, then ($\mathfrak{J}$ being defined as before) the birational correspondence $f^\tau \circ (f^\sigma)^{-1}$ between V^σ and V^τ must be everywhere biregular for all σ, τ in $\mathfrak{J}$. We will prove that this condition is also sufficient, at any rate if V is a k-open subset of a projective variety, defined over k. This will be an immediate consequence of the following result.

PROPOSITION 1. *Let k, k_0, $\mathfrak{J}$ be as in* (A). *Let V_0 be a variety, defined over k_0; let V be a projective (resp. affine) variety, defined over k; let f be a birational correspondence, defined over k, between V_0 and V. Then there is a projective (resp. affine) variety W and a birational correspondence F between V_0 and W, both defined over k_0, such that $F \circ f^{-1}$ is biregular at every point of V where the mappings $f^\sigma \circ f^{-1}$ are defined for all $\sigma \in \mathfrak{J}$.*

Let S be the ambient space of V, projective or affine; f may be regarded as a mapping of V_0 into S. Call $\sigma_1 = \epsilon, \sigma_2, \cdots, \sigma_n$ the elements of $\mathfrak{J}$, and put $F_1 = (f^{\sigma_1}, \cdots, f^{\sigma_n})$; this is a mapping of V_0 into the product $S \times \cdots \times S$ of n factors equal to S, and is defined over the compositum K of the fields k^σ. It is clear that $F_1 \circ f^{-1}$ is defined wherever all the $f^\sigma \circ f^{-1}$ are defined. Let x be a generic point of V_0 over k_0; let W_1 be the locus of $F_1(x)$ over K; put $u = F_1(x) = (x_1, \cdots, x_n)$. As $\sigma_1 = \epsilon$, we have $x_1 = f(x)$, so that the image of u by the mapping $f \circ F_1^{-1}$ is x_1; this shows that $f \circ F_1^{-1}$ is the mapping induced on W_1 by the projection of the product $S \times \cdots \times S$ onto its first factor, and is therefore everywhere defined. Thus the birational correspon-

dence $F_1 \circ f^{-1}$ between V and W_1 is biregular wherever all the $f^\sigma \circ f^{-1}$ are defined.

Let $z_1, \cdots, z_n$ be n points of S; if S is the projective m-space, put $z_i = (z_{i0}, \cdots, z_{im})$; and let z' be the point, in a projective space of suitable dimension, whose homogeneous coordinates are all the monomials $z_{1\mu_1} z_{2\mu_2} \cdots z_{n\mu_n}$, with $0 \leqq \mu_i \leqq m$ for every i. If S is the affine m-space, put $z_i = (z_{i1}, \cdots, z_{im})$, put $z_{i0} = 1$ for $1 \leqq i \leqq n$, and let z' be the point, in an affine space of suitable dimension, whose coordinates are the same monomials as before. In either case, put $z' = \Phi(z_1, \cdots, z_n)$; it is well-known that Φ is an everywhere biregular mapping of $S \times \cdots \times S$ onto its image in projective (resp. affine) space. Put now $F_2 = \Phi \circ F_1$; then F_2 is a birational correspondence between V_0 and $W_2 = \Phi(W_1)$, and $F_2 \circ f^{-1}$ is biregular wherever all the $f^\sigma \circ f^{-1}$ are defined.

If S is projective, let $(1, f_1(x), \cdots, f_m(x))$ be a set of homogeneous coordinates for $f(x)$; the f_μ are functions on V_0, defined over k. Put $f_0 = 1$. Then we have $F_2 = (g_0, \cdots, g_r)$, where the g_ρ are all the monomials

$$f_{\mu_1}^{\sigma_1} f_{\mu_2}^{\sigma_2} \cdots f_{\mu_n}^{\sigma_n}.$$

If ω is an automorphism of K over k, $g_\rho{}^\omega$ is again one of the g_ρ, which we may write as $g_{\omega(\rho)}$; the mapping $\rho \to \omega(\rho)$ determines a representation of Γ (the Galois group of K over k) as a group of permutations on the g_ρ. For a given ρ, let γ_ρ be the subgroup of Γ determined by $\omega(\rho) = \rho$; then, for $\omega \in \Gamma$, $\omega(\rho)$ takes a number of distinct values equal to the index d_ρ of γ_ρ in Γ. If K_ρ is the subfield of K consisting of the elements of K invariant under γ_ρ, g_ρ is defined over K_ρ; therefore, if $(\alpha_1, \cdots, \alpha_{d_\rho})$ is a basis of K_ρ over k_0, we may write $g_\rho = \sum_\nu \alpha_\nu h_{\rho\nu}$, where the $h_{\rho\nu}$, for $1 \leqq \nu \leqq d_\rho$, are functions on V_0, defined over k_0. Then we have, for all $\omega \in \Gamma$:

$$g_{\omega(\rho)} = \sum_\nu \alpha_\nu{}^\omega h_{\rho\nu}.$$

If, in this relation, we take for ω a set of representatives of the d_ρ cosets of γ_ρ in Γ, we get a linear substitution expressing the d_ρ distinct functions $g_{\omega(\rho)}$ in terms of the d_ρ functions $h_{\rho\nu}$; and, since K_ρ is separable over k_0, that substitution is invertible. From this it follows immediately that, if we call $F(x)$ the point whose homogeneous coordinates are all the functions $h_{\rho\nu}$ (where ρ runs through a set of representatives for the classes of equivalence determined by the permutation group Γ on the set $\{0, 1, \cdots, r\}$, and where, for each ρ, we take $1 \leqq \nu \leqq d_\rho$), F is of the form $\Psi \circ F_2$, where Ψ is an automorphism of the ambient projective space of W_2. If S is affine, we put

$f=(f_1,\cdots,f_m)$, $f_0=1$, and we define F by the same formulas as in the projective case, but regard it as a mapping of V_0 into an affine space; then we have again $F=\Psi\circ F_2$, Ψ being now an automorphism of the ambient affine space of W_2. In either case, the mapping F is defined over k_0; if W is the locus of $F(x)$ over k_0, W and F have the properties required by our proposition.

3. Before applying this to the problems (A′) and (B), we need a general lemma:

LEMMA 1. *Let f be a birational correspondence between two varieties U and V; let k be a field of definition for U, V, f. Then the sets of points where f and f^{-1} are respectively biregular are k-open, and f determines a k-isomorphism between them.*

Call U' the set of points of U where f is defined, U'' the set of points of U where it is biregular; call V' the set of points of V where f^{-1} is defined, V'' the set where it is biregular. By [4], App., Prop. 8, U' and V' are k-open. Call f' the restriction of f to $U'\times V'$ (i.e. the birational correspondence between U' and V' whose graph is the set-theoretic intersection of the graph of f with $U'\times V'$). If f is biregular at a point a of U, it is defined at a, so that $a\in U'$; and, if we put $b=f(a)$, f^{-1} is defined at b, so that $b\in V'$; therefore (a,b) is on the graph of f', and f' is defined at a. Conversely, let a be a point of U' where f' is defined; put $b=f'(a)$; then b is in V', so that f^{-1} is defined at b; as U' and V' are open, f is then defined at a, and we have $b=f(a)$; thus f is biregular at a. This shows that U'' is the set of points of U' where f' is defined; similarly V'' is the set of points of V' where f'^{-1} is defined; this implies that they are k-open. If f'' is the restriction of f to $U''\times V''$, it is everywhere biregular by definition (i.e., f'' is biregular at every point of U'', and f''^{-1} is so at every point of V'').

In order to formulate our results on problems (A′) and (B), we will say that a variety U, defined over a field k, is *projectively* (resp. *affinely*) *embeddable over k* if it is k-isomorphic to a k-open subset of a projective (resp. affine) variety, defined over k.

THEOREM 2. *Problem* (B) *has a solution (V_0',F') provided V is projectively embeddable over k and the birational correspondence $f^\sigma\circ f^{-1}$ between V and V^σ is everywhere biregular for every isomorphism σ of k over k_0 into $\bar{k}_0$. When that is so, V_0' is projectively embeddable over k_0; it is affinely embeddable over k_0 if V is so over k.*

We may assume V to be a k-open subset of a projective (resp. affine) variety, defined over k. Take W and F as in Proposition 1; then $F \circ f^{-1}$ is biregular at every point of V; therefore, by Lemma 1, it is a k-isomorphism between V and the k-open subset V_0' of W where $f \circ F^{-1}$ is biregular. As the $f^\sigma \circ f^{-1}$ are everywhere biregular V_0' is also the subset of W where $f^\sigma \circ F^{-1}$ is biregular, for every σ; therefore it is invariant under all automorphisms of k_0 over k_0, so that it is k_0-open, by [4], App., Prop. 9. Then, if F' is the restriction of F to $V_0 \times V_0'$, (V_0', F') is a solution of (B).

THEOREM 3. *Problem* (A′) *has a solution, i.e., problem* (A) *has a solution* (V_0, f) *for which* f *is everywhere biregular, provided* V *is projectively embeddable over* k *and the* $f_{\tau,\sigma}$ *are everywhere biregular and satisfy the conditions in Theorem* 1. *When that is so,* V_0 *is projectively embeddable over* k_0; *it is affinely embeddable over* k_0 *if* V *is so over* k. *The solution is unique up to a* k_0*-isomorphism.*

Section II. Regular Extensions of the Groundfield.

4. Let now k denote the groundfield. Let T be a variety, defined over k; let t be a generic point of T over k. When we denote by V_t a variety, defined over $k(t)$, we will agree, whenever t' is also a generic point of T over k, to denote by $V_{t'}$ the transform of V_t by the isomorphism of $k(t)$ onto $k(t')$ over k which maps t onto t'. Similarly, if a mapping, defined over $k(t)$, is denoted by f_t, $f_{t'}$ will denote its transform by the same isomorphism; if t, t', t'' are three independent generic points of T over k, and $f_{t',t}$ is a mapping, defined over $k(t,t')$, we denote by $f_{t'',t'}$ the transform of $f_{t',t}$ by the isomorphism of $k(t,t')$ onto $k(t',t'')$ over k which maps (t,t') onto (t',t''); etc.

Let V_t be a variety, defined over $k(t)$; assume that there is a variety V, defined over k, and a birational correspondence f_t, defined over $k(t)$, between V and V_t; then $f_{t'} \circ f_t^{-1}$ is a birational correspondence between V_t and $V_{t'}$. We therefore modify problem (P) of the introduction as follows:

(C) *Let* T *be a variety, defined over a field* k; *let* t, t' *be independent generic points of* T *over* k. *Let* V_t *be a variety, defined over* $k(t)$; *let* $f_{t',t}$ *be a birational correspondence, defined over* $k(t,t')$, *between* V_t *and* $V_{t'}$. *Find a variety* V, *defined over* k, *and a birational correspondence* f_t, *defined over* $k(t)$, *between* V *and* V_t, *such that* $f_{t',t} = f_{t'} \circ f_t^{-1}$.

THEOREM 4. *Problem* (C) *has a solution if and only if* $f_{t',t}$ *satisfies the condition:*

(i) $$f_{t'',t} = f_{t'',t'} \circ f_{t',t}$$

where t'' is a generic point of T over $k(t,t')$. When that is so, the solution is unique, up to a birational transformation on V, defined over k.

The condition is obviously necessary. The proof for the unicity of the solution, when one exists, is quite similar to the proof of the corresponding statement in Theorem 1. Now, assuming (i) to be fulfilled, we shall construct a solution of (C). We may replace T by any birational transform of T over k, and so we may assume that T is an affine variety. Similarly we may assume that V_t is an affine variety; and, taking x to be a generic point of V_t over $k(t)$, we may replace x by (x,t) and V_t by the locus of (x,t) over $k(t)$; after that is done, V_t is still an affine variety, and we have $k(t) \subset k(x)$; from now on, assume that this is so, and assume that x has been taken generic on V_t over $k(t,t',t'')$. By [4], App., Prop. 1, $k(x)$ is a regular extension of k; call X the locus of x over k. Put $x' = f_{t',t}(x)$; this is a generic point of $V_{t'}$ over $k(t,t',t'')$; by the definition of $V_{t'}$, this implies that there is an isomorphism of $k(t,x)$ onto $k(t',x')$ over k, mapping t onto t' and x onto x'; therefore we have $k(t') \subset k(x')$, hence $k(x,t') \subset k(x,x')$. As the definition of x' shows $k(x,x')$ to be contained in $k(t',t,x)$, i.e. in $k(x,t')$, it follows that we have $k(x,x') = k(x,t')$; therefore x' has a locus W_x over $k(x)$. Let $k(v)$ be the smallest field of definition containing k for W_x; as $k(v) \subset k(x)$, $k(v)$ is a regular extension of k. Call V the locus of v over k; we may write $v = G(x)$, where G is a mapping of X into V, defined over k.

If we put $x'' = f_{t'',t}(x)$, W_x is also the locus of x'' over $k(x)$; as the fields $k(x,x')$ and $k(x,x'')$ are respectively the same as $k(x,t')$ and $k(x,t'')$ and are therefore algebraically independent over $k(x)$, W_x is also the locus of x'' over $k(x,x')$. But (i) may be written $x'' = f_{t'',t'}(x')$; therefore $W_{x'}$ is the same as W_x. This implies that the isomorphism of $k(x)$ onto $k(x')$ over k which maps x onto x' leaves invariant all the elements of the smallest field of definition of W_x, hence also all the elements of $k(v)$, so that we have $G(x) = G(x')$.

On the other hand, let K be an overfield of k, algebraically independent from $k(x,x')$ over k; if ϕ is any function on X, defined over K, it will induce on W_x a function which is defined over $K(v)$; if $\phi(x) = \phi(x')$, that function is a constant, so that its constant value must be in $K(v)$. This shows that $K(v)$ is the subfield of $K(x)$ consisting of the elements of $K(x)$ which are invariant under the isomorphism of $K(x)$ onto $K(x')$ over K mapping x onto x'.

Now the relation $x'' = f_{t'',t}(x)$ shows that x'' is rational over $k(t'',t,x)$,

i.e. over $k(t'', x)$, so that we may write $x'' = \phi_{t''}(x)$, where $\phi_{t''}$ is a mapping of X into $V_{t''}$, defined over $k(t'')$. The relation $x'' = f_{t'',t'}(x')$ may then be written as $x'' = \phi_{t''}(x')$. Applying to the field $K = k(t'')$ and to the function $\phi_{t''}$ what we have proved above, we conclude from this that $k(x'') \subset k(t'', v)$. As we have $G(x) = G(x')$, hence also $G(x) = G(x'')$, the isomorphism of $k(x)$ onto $k(x'')$ over k which maps x onto x'' leaves $v = G(x)$ invariant; applying the inverse of that isomorphism to the relation $k(x'') \subset k(t'', v)$, we get $k(x) \subset k(t, v)$, hence $k(x) = k(t, v)$ since $k(t)$ and $k(v)$ are both contained in $k(x)$. Also, since $k(x)$ and $k(t')$ are algebraically independent over k, the same is true of $k(v)$ and $k(t')$; as the isomorphism of $k(x')$ onto $k(x)$ over k which maps x' onto x maps t' onto t and v onto itself, this implies that $k(v)$ and $k(t)$ are algebraically independent over k. As the relation $k(x) = k(t, v)$ can also be written $k(t, x) = k(t, v)$, we conclude that V_t and V are birationally equivalent over $k(t)$, so that we may write $x = f_t(v)$, where f_t is a birational correspondence between V and V_t, defined over $k(t)$. Then we have $x' = f_{t'}(v)$. Therefore (V, f_t) is a solution of our problem. We also see that X is birationally equivalent to $T \times V$ over k.

5. Just as in Section I, we consider the problem (C′) which consists in finding a solution (V, f_t) of (C) such that f_t is everywhere biregular. For such a solution to exist, it is necessary that $f_{t',t}$ should be everywhere biregular; it will be shown that this is sufficient.

As in Section I, if we make use of Theorem 4, we see that (C′) may be reformulated as follows:

(D) *Let k, T and t be as in* (C); *let V and V_t be varieties, respectively defined over k and over $k(t)$; let f_t be a birational correspondence, defined over $k(t)$, between V and V_t. Find a variety V' and a birational correspondence F between V and V', both defined over k, such that the birational correspondence $f_t \circ F^{-1}$ between V' and V_t is everywhere biregular.*

In order to solve (D), we need some preliminary results.

LEMMA 2. *Let F and H be mappings of a variety X into two varieties W, T, all these being defined over a field k; x being a generic point of X over k, assume that $t = H(x)$ is generic over k on T and that x has a locus V_t over $k(t)$. Let F_t be the mapping of V_t into W induced by F on V_t. Then F is defined at every point of V_t where F_t and H are both defined.*

It is clearly enough to treat the case in which X is an affine variety and W is the affine line. Then F_t is the function on V_t, defined over $k(t)$,

such that $F_t(x) = F(x)$. If F_t is defined at a point a of V_t, we can write it as $F_t(x) = P(x)/Q(x)$, where P, Q are polynomials with coefficients in $k(t)$, such that $Q(a) \neq 0$. More explicitly, we have

$$P(X) = \sum_i \lambda_i(t) M_i(X), \qquad Q(X) = \sum_j \mu_j(t) N_j(X),$$

where the λ_i, μ_j are functions on T, defined over k, and the M_i, N_j are monomials in the indeterminates (X); and we have

$$\sum_j \mu_j(t) N_j(a) \neq 0. \tag{1}$$

Then we have $F(x) = \Phi(x)/\Psi(x)$, with

$$\Phi(x) = \sum_i \lambda_i(H(x)) M_i(x), \qquad \Psi(x) = \sum_j \mu_j(H(x)) N_j(x). \tag{2}$$

As a is on V_t, (t, a) is a specialization of (t, x) over k; if H is defined at a, we must have $H(a) = t$. As t is generic on T over k, the functions λ_i, μ_j are defined at t; therefore the functions $\lambda_i \circ H$, $\mu_j \circ H$ are defined at a on X, with the values $\lambda_i(t)$, $\mu_j(t)$. That being so, the relations (1), (2) show that F is defined at a on V.

PROPOSITION 2. *Let k, T, t, t' be as in* (C); *let V be a variety, defined over k; let V_t be a variety, defined and projectively (resp. affinely) embeddable over $k(t)$; let f_t be a birational correspondence, defined over $k(t)$, between V and V_t. Then:*

(i) *if a is a point of V_t where $f_{t'} \circ f_t^{-1}$ is biregular, there is an affine variety W and a birational correspondence F between V and W, both defined over k, such that $F \circ f_t^{-1}$ is biregular at a;*

(ii) *if $f_{t'} \circ f_t^{-1}$ is everywhere biregular, there is a variety W, defined and projectively (resp. affinely) embeddable over k, and a birational correspondence F between V and W, defined over k, such that $F \circ f_t^{-1}$ is everywhere biregular.*

We may assume that V_t is a $k(t)$-open subset of a variety, defined over $k(t)$, in a projective (resp. affine) space S. We may also assume that T is a projective (resp. affine) variety; let S' be its ambient space. If S, S' are affine, $S \times S'$ is an affine space; if they are projective, call Φ the well-known biregular embedding of $S \times S'$ into a projective space S'' of suitable dimension. Let v be generic on V over $k(t, t')$, and put $x = f_t(v)$. We may replace V_t by a suitable $k(t)$-open subset of the locus of (x, t) over $k(t)$ in the affine case, of $\Phi(x, t)$ over $k(t)$ in the projective case; after that is done,

we have $k(t) \subset k(x)$, and therefore $k(x) = k(t, x) = k(t, v)$, so that x has a locus X over k, birationally equivalent to $T \times V$, and that we may write $t = H(x)$, where H is a mapping of X into T, defined over k; moreover, the mapping H is everywhere defined on X.

Now, since X is birationally equivalent to $T \times V$ over k, and V is birationally equivalent to V_t over $k(t)$, X is birationally equivalent to $T \times V_t$ over $k(t)$. More explicitly, if we put $x' = f_{t'}(v)$, x' is generic over $k(t)$ on X, and we have $k(x') = k(t', v)$, hence $k(t, x') = k(t, t', x)$, so that we may write $x' = g_t(t', x)$, where g_t is a birational correspondence, defined over $k(t)$, between $T \times V_t$ and X. We have $t' = H(x')$, and we may write $x = \phi_t(x')$, where ϕ_t is a mapping of X into V_t, defined over $k(t)$; then (H, ϕ_t) is the mapping of X into $T \times V_t$, inverse to g_t. The mapping g_t induces on the subvariety $t' \times V_t$ of $T \times V_t$ the mapping $(t', x) \to x' = f_{t'}(f_t^{-1}(x))$; and ϕ_t induces on $V_{t'}$ the mapping $x' \to x$, i. e. the mapping $f_t \circ f_{t'}^{-1}$. Applying Lemma 2, we see that g_t is defined at (t', a) whenever a is a point of V_t where $f_{t'} \circ f_t^{-1}$ is defined, and that ϕ_t is defined at every point of $V_{t'}$ where $f_t \circ f_{t'}^{-1}$ is defined. Therefore g_t is biregular at (t', a) whenever a is a point of V_t where $f_{t'} \circ f_t^{-1}$ is biregular.

Now let A_0 be the $k(t)$-closed subset of $T \times V_t$ where g_t is not biregular; and assume first that a is a point of V_t with the property stated in (i). Then (t', a) is not in A_0, so that $T \times a$ is not contained in A_0; let A_1 be the (non-dense) $k(t, a)$-closed subset of T consisting of those points t_1 such that $(t_1, a) \in A_0$. By [4], App., Prop. 12, there is a $\bar{k}$-closed subset A_2 of T containing all $\bar{k}$-closed subsets of T contained in A_1; in particular, every point of A_1 which is algebraic over k must be in A_2. Let A_3 be the union of the components of A_2 and of their conjugates over k; put $T' = T - A_3$; this is a k-open subset of T such that, if t_1 is any algebraic point over k in T', g_t is biregular at (t_1, a).

On the other hand, assume, as in (ii), that $f_{t'} \circ f_t^{-1}$ is everywhere biregular. Then g_t is biregular at every point of $t' \times V_t$, so that A_0 has no point in common with $t' \times V_t$. This implies that the projection of A_0 on T is non-dense in T, so that, if we call A_1' the closure of that projection, it is a (non-dense) $k(t)$-closed subset of T. Let A_2' be the maximal $\bar{k}$-closed subset of T contained in A_1'; let A_3' be the union of the components of A_2' and of their conjugates over k; put $T'' = T - A_3'$. Then T'' is k-open on T; and, if t_1 is any algebraic point over k in T'', g_t is biregular at every point of $t_1 \times V_t$.

Now let t_1 be a separably algebraic point over k in T' (resp. T''); if k is finite, we may take for t_1 any algebraic point over k in T' (resp. T''),

since in that case every algebraic extension of k is separable; if k is infinite, we apply [4], App., Prop. 13. Let $t_1, \cdots, t_n$ be the distinct conjugates of t_1 over k. As they are in T' (resp. T''), g_t is biregular at (t_i, a) (resp. at every point of $t_i \times V_t$), and a fortiori at (t_i, x), for $1 \leqq i \leqq n$; therefore it induces on $t_i \times V_t$ a birational correspondence g_i between V_t and the locus V_i of the point $g_i(x) = g_t(t_i, x)$ over $k(t, t_i)$ in the projective (resp. affine) ambient space of X; and g_i is biregular at a (resp. at every point of V_t). But, as we have already observed, the relation $k(x) = k(t, v)$ shows that X is birationally equivalent to $T \times V$; we may write $x = f(t, v)$, where f is a birational correspondence between $T \times V$ and X, defined over k; then we have $x' = f(t', v)$; and f is the product of g_t and of the birational correspondence $(t', v) \to (t', x)$ between $T \times V$ and $T \times V_t$. As the latter correspondence is biregular at (t_i, v), and g_t is biregular at (t_i, x), for $1 \leqq i \leqq n$, we see that f is biregular at (t_i, v), and that we have

$$g_i(x) = g_t(t_i, x) = f(t_i, v).$$

As the point $f(t_i, v)$ has the same locus over $k(t_i)$ as over $k(t, t_i)$, this shows that V_i is defined over $k(t_i)$. As every automorphism of $\bar{k}$ over k can be extended to an automorphism of $\bar{k}(v)$ over $k(v)$, this also shows that V_i is the transform of V_1 by the isomorphism of $k(t_1)$ onto $k(t_i)$ over k which maps t_1 onto t_i. Also, if f_i is the mapping of V into V_i, defined over $k(t_i)$, which is such that $f_i(v) = f(t_i, v)$, we have $f_i = g_i \circ f_t$; and f_i is the transform of f_1 by the isomorphism of $k(t_1)$ onto $k(t_i)$ over k which maps t_1 onto t_i.

Now apply Proposition 1 to the variety V, defined over the groundfield k, to the variety V_1, defined over $k(t_1)$, and to the birational correspondence f_1; this gives a projective (resp. affine) variety W and a birational correspondence F between V and W, both defined over k, such that $F \circ f_1^{-1}$ is biregular wherever all the $f_i \circ f_1^{-1}$ are defined, i.e. wherever all the $g_i \circ g_1^{-1}$ are defined. Now, in case (i), all the g_i are biregular at a, so that all the $g_i \circ g_1^{-1}$ are biregular at the point $g_1(a)$; therefore $F \circ f_t^{-1}$, which is the same as $(F \circ f_1^{-1}) \circ g_1$, is biregular at a; as this involves merely a local property of W at the image of a by that mapping, we may replace W, in the projective case, by one of its affine representatives. Thus we have solved our problem in case (i). In case (ii), g_i is biregular at every point of V_t; as we have just shown, this implies that $F \circ f_t^{-1}$ is biregular at every point of V_t, so that it determines an isomorphism of V_t onto a $k(t)$-open subset W' of W. The assumption in (ii) implies that W' is invariant under the isomorphism of $k(t)$ onto $k(t')$ over k which maps t onto t'. From this and from [4], App., Prop. 9, it follows easily that W' is k-open; thus (W', F) is a solution of our problem.

COROLLARY. *Let k, T, t and t' be as in* (C); *let V be a variety, defined over k; let V_t be a variety, defined over $k(t)$; let f_t be a birational correspondence between V and V_t, defined over $k(t)$ and such that $f_{t'} \circ f_t^{-1}$ is everywhere biregular. Then, if a is any point of V_t, there is an affine variety W and a birational correspondence F between V and W, both defined over k, such that $F \circ f_t^{-1}$ is biregular at a.*

We may assume that t' has been taken generic on T over $k(t,a)$; take t'' generic on T over $k(t,t',a)$. Call a', a'' the images of a by $f_{t'} \circ f_t^{-1}$ and by $f_{t''} \circ f_t^{-1}$, respectively. The isomorphism of $k(t,a,t')$ onto $k(t,a,t'')$ over $k(t,a)$ which maps t' onto t'' maps a' onto a''; therefore, if $V_{t'a}$ is a representative of the (abstract) variety $V_{t'}$ on which a' has a representative a_a', the point a'' of $V_{t''}$ has a representative a_a'' on $V_{t''a}$. Let $f_{t'a}$ be the birational correspondence between V and $V_{t'a}$ which is determined by $f_{t'}$. As $f_{t''} \circ f_{t'}^{-1}$ is everywhere biregular and maps a' onto a'', $f_{t''a} \circ f_{t'a}^{-1}$ is biregular at a_a'. Applying Proposition 2(i) to V, $V_{t'a}$ and $f_{t'a}$, we get a solution (W, F) of our problem.

6. Now we can deal with problems (D) and (C').

THEOREM 5. *Problem* (D) *has a solution if and only if $f_{t'} \circ f_t^{-1}$ is everywhere biregular for t' generic over $k(t)$ on T.*

The condition being obviously necessary, assume that it is fulfilled. By the corollary of Proposition 2, there is, to every point a of V_t, an affine variety W_a and a birational correspondence F_a between V and W_a, both defined over k, such that $F_a \circ f_t^{-1}$ is biregular at a; call Ω_a the $k(t)$-open subset of V_t where $F_a \circ f_t^{-1}$ is biregular, and call W_a' its image on W_a by $F_a \circ f_t^{-1}$, which is a $k(t)$-open subset of W_a. Then W_a' is the subset of W_a where $f_t \circ F_a^{-1}$ is biregular; as in the proof of Proposition 2, this implies that W_a' is invariant under the isomorphism of $k(t)$ onto $k(t')$ over k which maps t onto t', and we again conclude from this that W_a' is k-open. As we have $a \in \Omega_a$ for every $a \in V_t$, the open sets Ω_a form a covering of V_t; by the well-known "compactoid" property of open sets in the Zariski topology, there must be finitely many points a_α on V such that the sets Ω_{a_α} cover V_t. Then the k-open subsets W_{a_α}' of the affine varieties W_{a_α}, together with the birational correspondences $F_{a_\beta} \circ F_{a_\alpha}^{-1}$ between them, define an abstract variety, which, together with the obvious birational correspondence between it and V, solves our problem.

THEOREM 6. *Problem* (C') *has a solution, i.e., problem* (C) *has a*

solution (V, f_t) *for which* f_t *is everywhere biregular, if and only if* $f_{V,t}$ *is everywhere biregular and satisfies condition* (i) *in Theorem* 4. *The solution is unique up to a* k*-isomorphism.*

This is an immediate consequence of Theorems 4 and 5.

7. As to the projective or affine embeddability of the solution of problems (D) and (C′), we have the following result.

THEOREM 7. *Let* V *be a variety, defined over a field* k*, and projectively (resp. affinely) embeddable over an overfield* K *of* k. *Then* V *is projectively (resp. affinely) embeddable over* k *provided* (i) K *is separable over* k *or* (ii) V *is everywhere normal with reference to* k.

The assumption means that there is a birational correspondence f, defined over K and biregular at every point of V, between V and a subvariety of a projective (resp. affine) space; if we regard f as a mapping of V into that space, it has a smallest field of definition k' containing k; we may replace K by k'; after that is done, K is finitely generated over k. If K is separable over k, it is a regular extension $k_1(t)$ of the algebraic closure k_1 of k in K, and k_1 is a separably algebraic extension of k of finite degree. Proposition 2(ii) shows that V is then projectively (resp. affinely) embeddable over k_1; by Theorem 2, this implies that the same is true over k; this completes the proof in case (i). If K is not separable over k, let k^* be the union of the fields $k^{p^{-n}}$, for $n = 1, 2, \cdots$; then the compositum K^* of K and k^* is separable over k^*, so that, by what we have just proved, V is projectively (resp. affinely) embeddable over k^*. In order to deal with case (ii), it is therefore enough to prove our theorem in the case in which V is everywhere normal with reference to k, and K is purely inseparable over k; I owe the proof for this to T. Matsusaka; it is as follows.

We may again assume that K is finitely generated over k; as it is purely inseparable, it is contained in some field $k' = k^{1/q}$, where q is a power of the characteristic. Then there is a mapping f' of V into a projective (resp. affine) space, defined over k', such that f' determines a birational correspondence, biregular at every point of V, between V and the closure W' of its image by f'; then W' is a projective (resp. affine) variety, defined over k', and f' determines a k'-isomorphism between V and a k'-open subset of W'. Call π the automorphism $\xi \to \xi^q$ of the universal domain; put $W = W'^{\pi}$; W is then a projective (resp. affine) variety, defined over k. Let x be a generic point of V over k; then W' is the locus of the point $y' = f'(x)$ over k', and W is

the locus of the point $y=y'^{\pi}$ over k. As y' is rational over $k'(x)$, y is so over $k(x^{\pi})$; we may write $y=g(x)$, where g is a mapping of V into W, defined over k; as we have $k'(y')=k'(x)$, we have $k(y)=k(x^{\pi})$, which implies that $k(x)$ is purely inseparable over $k(y)$. In the projective case, let U be the projective variety derived from W by normalization in the field $k(x)$[1]; U is birationally equivalent to V over k; let z be the point of U which corresponds to x on V. In the affine case, we take for z a point in a suitable affine space such that $k[z]$ is the integral closure of the ring $k[y]$ in the field $k(x)$, and for U the locus of z over k. In either case we may write $z=f(x)$, where f is a birational correspondence between V and U, defined over k. By definition, U is everywhere normal with reference to k, and the mapping $h=g\circ f^{-1}$ of U into W is everywhere defined and such that the (set-theoretic) inverse image of every point of W for that mapping consists of finitely many points of U. Let a be any point of V; let (a,b) be a specialization of (x,z) over $x\to a$ with reference to k; then, as h is defined at b, $(a,b,h(b))$ is a specialization of (x,z,y) over k. As f' is defined at a, g is also defined there, so that we must have $h(b)=g(a)$; therefore b is one of the finitely many points of U whose image by h is $g(a)$. As V is normal at a by assumption, with reference to k, this implies that f is defined at a, and that we have $b=f(a)$. We have $g(a)=f'(a)^{\pi}$, hence $f'(a)=g(a)^{\pi^{-1}}$; as $g(a)=h(b)$, this shows that $f'(a)$ is the unique specialization of y' over $z\to b$ with reference to k'; as f' is biregular at a, f'^{-1} is defined at $f'(a)$, and therefore x has no other specialization than a over $z\to b$ with reference to k', hence also with reference to k by F-II$_1$, Prop. 3. As U is normal at b, with reference to k, this implies that f^{-1} is defined at $b=f(a)$. We have thus shown that f is biregular at every point of V, so that it is a k-isomorphism between V and a k-open subset of U.

As a special case (already contained in Proposition 2), we see that, in problem (D), V' is projectively (resp. affinely) embeddable over k if V_t is so over $k(t)$; similarly, in problem (C'), V is projectively (resp. affinely) embeddable over k if V_t is so over $k(t)$.

8. In [4], the construction carried out in Nos. 7-9 can be advantageously replaced by the application of our Theorem 6 to the situation described in No. 6 of that paper. The application is entirely straightforward, so that no further details need be given; this shows that the recourse to the Lang-Weil

[1] U is the "derived normal model of W in the field $k(x)$" according to Zariski's definition ([5], pp. 69-70); cf. also [3].

Theorem, i.e., in substance, to the so-called "Riemann hypothesis" in the case of a finite groundfield (loc. cit., p. 374) was unnecessary; so is the assumption of normality in the final result (loc. cit., p. 375); normality had to be assumed there merely because of the use made of the Chow point in the construction on p. 370, whereas in the present paper a different device was adopted (in the proof of Proposition 1). Of course, in the main theorem of [4] (p. 375), parts (i) and (ii) remain unchanged. For the sake of completeness, we give here the improved result by which part (iii) of that theorem may now be replaced:

PROPOSITION 3. *Let G be a group and W a chunk of transformation-space with respect to G, both defined over k. Then there is a transformation-space S with respect to G, and a birational correspondence f between W and S, both defined over k, with the following properties:* (a) *f is biregular at every point of W*; (b) *for every $s \in G$ and $a \in W$ such that sa is defined, we have $f(sa) = sf(a)$*; (c) *every point of S can be written in the form $sf(a)$, with $s \in G$ and $a \in W$. Moreover, S is uniquely determined by these properties up to a k-isomorphism compatible with the operations of G.*

UNIVERSITY OF CHICAGO.

REFERENCES.

[1] W. L. Chow, "Abelian varieties over function-fields," *Transactions of the American Mathematical Society*, vol. 78 (1955), pp. 253-275.

[2] S. Lang, "Abelian varieties over finite fields," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 174-176.

[3] T. Matsusaka, "A note on my paper 'Some theorems on abelian varieties'," *Nat. Sc. Rep. Ochanomizu Univ.*, vol. 5 (1954), pp. 21-23.

[4] A. Weil, "On algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955), pp. 355-391.

[5] O. Zariski, "Theory and applications of holomorphic functions on algebraic varieties over arbitrary groundfields," *Memoirs of the American Mathematical Society*, no. 5 (1951), pp. 1-90.

ON THE REGULARITY REGIONS OF THE SOLUTIONS OF THE PARTIAL DIFFERENTIAL EQUATIONS OF CAUCHY-KOWALEWSKY.*

By AUREL WINTNER.

---

1. Let $F = F(z, w, s, v)$ be an analytic functions which is regular in a neighborhood of the point $(0, 0, 0, 0)$ of the complex (z, w, s, v)-space and, in a neighborhood of the point $(0, 0)$ of the (z, w)-space, let $s = s(z, w)$ be the (unique, regular) solution of

$$(I) \qquad s_z = F(z, w, s, s_w), \qquad s(0, w) \equiv 0$$

(Cauchy-Kowalewsky). If $F(z, w, s, v)$ is of the particular form $F(z, s)$, then (I) and its solution $s = s(z, w)$ will simplify to

$$(II) \qquad ds/dz = F(z, s), \qquad s(0) \equiv 0$$

and $s = s(z)$, respectively. A question concerning an absolute constant in the problem of a lower bound for the convergence radius of the power series $s(z)$ in the special (ordinary) case (II) of the (partial) differential equation (I) was dealt with in [12]. The corresponding question for associated radii of regularity of the solution $s(z, w)$ of the general partial differential equation (I) is much more complex. The only absolute lower bounds known for the associated radii of the double power series $s(z, w)$ are those obtained by Perron [9]. His result will be improved below to some extent, and so as to imply an affirmative answer to one of the questions of function-theoretical interest (concerning *small* functions F of the four variables z, w, s, v) but the problem of the "best absolute constants" remains unsolved.

It appears therefore natural that what should be settled first is the problem of the "best" associated radii of the solution $s(z, w)$ of the initial value problem of a quasi-linear partial differential equation, say of

$$(III) \qquad s_z = f(z, w, s) + g(z, w, s)s_w, \qquad s(0, w) \equiv 0,$$

with two functions f, g of (z, w, s) which are regular in a neighborhood of

---

* Received February 29, 1956.

the point $(z, w, s) = (0, 0, 0)$. In fact, the problem for (III) is intermediary between the solved problem for (II) (where $g = 0$) and the unsolved problem for (I) (where F can be more general than $F(z, w, s, v) = f + gv$; $v = s_w$).

2. In order to deal with problems of "best associated radii" in the analytic problem, it seemed to be natural to consult first the literature of the corresponding problem in the real domain. But what the literature contains in this regard proved to be of no avail. It is true true Kamke's book of 1930 contains a theorem which implies, among other things, the following assertion (Satz 4, § 173, pp. 335-336 in [3]): If $f(x, y, u)$ and $g(x, y, u)$ are real-valued, uniformly continuous functions on a parallelepipedon

$$0 \leqq x < a, \qquad -b < y < b, \qquad -c < u < c \tag{IV}$$

and if both functions have, with respect to y and u, continuous first derivatives satisfying

$$|\, d(x, y, u)\,| \leqq A = \text{const. on (IV), where } d = f_y, f_u, g_y, \tag{V}$$

then there exists an absolute constant $\Theta > 0$ (in fact, some

$$\Theta \geqq \tfrac{1}{2} \log 3; \tag{VI}$$

cf. [3], p. 336, footnote) in such a way that the initial value problem

$$u_x = f(x, y, u)\, u_y + g(x, y, u), \qquad u(0, y) \equiv 0, \tag{VII}$$

which is the analogue of (III) in the real field, has on the rectangle

$$0 \leqq x < \min\,(a, \Theta/A), \qquad -b < y < b \tag{VIII}$$

a (unique) continuously differentiable solution $u = u(x, y)$. But Kamke's proof is erroneous (and, as a matter of fact, the assertion is false for *any* absolute $\Theta > 0$). The trouble is at the very beginning of the proof ([3], p. 336), where, on the one hand, $f(x, y, u)$ and $g(x, y, u)$ are extended from the parallelepipedon (IV), where $0 < b < \infty$, to the infinite slab which is the case $b = \infty$ of (IV) and, on the other hand, the tacit (but erroneous) assumption is made that, when the point (x, y) is in the rectangle (VIII), the point (x, y, u) belonging to the solution $u = u(x, y)$ will stay in the parallelepipedon (IV) belonging to the *given* $b < \infty$. The mistake was pointed out by Kamke himself in a *Nachtrag* to the second edition (1944) of his book (1930) and the result is stated correctly in [4], p. 41, by assuming that $b = \infty$ in (IV) (but the last footnote on p. 42 of [4] is misleading, since the coefficient $f(x, y, u)$ of (VII) above is not allowed to

contain u in [9]). Under the assumption $b = \infty$, following a suggestion of Ważewski (cf. [8], p. 2), Perausówna [8] determined the best value of the absolute constant Θ, by showing that (VI) can be improved to $\Theta \geqq 1$ and that the assumption $\Theta < 1$ is disproved by a simple example; so that $\Theta = 1$ (concerning quasi-linear *systems*, cf. [11]).

It is easy to see that Kamke's proof (with $b = \infty$) and Perausówna's improvement of it (again with $b = \infty$) apply in the complex field also. But what then results applies only to a very particular case of the problem of (III). In fact, Liouville's theorem shows that the analogues of (V) and of the case $b = \infty$ of (IV) in the complex field compel the coefficients of (III) to be of the form

$$f(z, w, s) = \alpha(z)w + \beta(z)s, \qquad g(z, w, s) = \gamma(z)w + \delta(z)s.$$

3. In what follows, the problem of "best associated radii" of (III) will be settled (Theorem 1) by an adaptation of the method used by Perron [10], pp. 557-562, in the particular case $f(x, y, u) = f(x, y)$ of the real equation (VII). By using a device contained already in Kowalewsky's thesis [5] (where the case of her § II is reduced to the case of her § I), the result on the quasi-linear problem (III) will lead to results on the general problem (I) also (Section 7). What thus results for the general case (I) is substantially finer than are the known estimates of the associate radii (even though one cannot readily speak of "best" constants in the general case).

The method, being that of the successive approximations, is such as to supply for (I) a result (Theorem 3) which does not concern the absolute constants but is of independent interest.

4. The following theorem (which, in view of Section 3, contains the central fact of the theory) will be proved first.

THEOREM 1. *Let a, b, c, L, M be five positive numbers, z, w, s three complex variables, finally*

$$f(z, w, s), \qquad g(z, w, s) \tag{1}$$

two functions which are regular on the (z, w, s)-domain

$$|z| < a, \qquad |w| + L|z| < b, \qquad |s| < c \tag{2}$$

and satisfy the inequalities

$$|f| < L, \quad |g| < M \text{ on } (2). \tag{3}$$

Then that solution

$$s = s(z, w) \tag{4}$$

(*Cauchy-Kowalewsky*) *of the quasi-linear partial differential equation*

$$s_z = f(z, w, s)s_w + g(z, w, s) \tag{5}$$

which belongs to the initial condition

$$s(0, w) \equiv 0 \tag{6}$$

is regular on the (z, w)*-domain*

$$|z| < \min(a, c/M), \qquad |w| + L|z| < b \tag{7}$$

(*at least*; but the domain (7) cannot be improved in terms of absolute constants); *moreover, the solution* (4) *is subject to the inequality*

$$|s(z, w)| < c \text{ on } (7). \tag{8}$$

The proof of this theorem will be reduced to a particular case of it:

LEMMA. *Theorem* 1 *is true if* f (*though not necessarily* g) *is independent of* s, *that is, if* $f = f(z, w)$.

If x, w, s and f, g are *real*, then, under appropriate conditions of differentiability (but without the assumption of analyticity, hence, without assuming the applicability of the Cauchy-Kowalewsky theorem), the preceding Lemma is contained in Satz 1 of Perron [10], p. 557, which, in the real field, even generalizes the case $f(z, w, s) = f(z, w)$ of (5) to certain *systems* of partial differential equations. Perron's proof (1927), which was the starting point of *all* the recent developments concerning quasi-linear hyperbolic systems of partial differential equations in two independent variables (for a list of references, cf. [6], pp. 257-258), is based on an application of the process of sucessive approximations. Correspondingly, a glance at Perron's proof shows that it remains valid in the complex field, in the form stated by the Lemma above. As a matter of fact, the proof is now shorter than in Perron's non-analytic case, since only the first, and comparatively easier, half of the proof (pp. 557-559, ending with formula (38)) is needed; the second half of the proof (pp. 559-562), that dealing with the convergence of the *derivatives*, now becomes superfluous, since the uniform convergence of the functions implies that of the derivatives in the analytic case.

5. *Proof of Theorem* 1. For every positive integer n, consider the solution, say

$$(4_n) \qquad s^n = s^n(z, w)$$

(where n is not an exponent), of the partial differential equation

$$(5_n) \qquad s_z{}^n = f^n(z, w)\, s_w{}^n + g(z, w, s^n)$$

and of the initial condition

$$(6_n) \qquad s^n(0, w) \equiv 0,$$

where, if f and g are the two functions given in Theorem 1, the g of (5_n) is the given g, whereas the f^n of (5_n) is defined as follows:

$$(9_n) \qquad f^n(z, w) = f(z, w, s^{n-1}(z, w)).$$

This is an inductive definition, which can be carried out by starting at $n = 0$ with

$$(10) \qquad s^0 = s^0(z, w) \equiv 0.$$

For then, if the (local) theorem of Cauchy-Kowalewsky is applied to (5_n)-(6_n), with (9_n) first for $n = 1$, then for $n = 2, \cdots$, it is clear that each of the functions (4_n) is defined, as a unique regular function, on some neighborhood of the point $(0, 0)$ of the complex (z, w)-space. What fails to follow in this manner is that this neighborhood of $(0, 0)$ can be chosen independent of n. But it will now be concluded from the Lemma that the (z, w)-domain specified by (7), a domain which is independent of n, is such a neighborhood for every n, and that

$$(8_n) \qquad |s^n(z, w)| < c \text{ on } (7).$$

Suppose that, for a fixed n, the function element $s^{n-1}(z, w)$ is known to be regular on (7), and that (8_{n-1}) is true for this n. It will be sufficient to show that both of these assumptions remain valid if $n - 1$ is replaced by n. In view of (10), the induction hypotheses are satisfied if $n - 1 = 0$. If $n - 1$ is fixed (and positive), then the regularity of $s^{n-1}(z, w)$ on (7) and the assumption (8_{n-1}) imply that, since the functions (1) are regular on (2) and satisfy (3), it follows from (9_n) that both functions

$$(1_n) \qquad f^n(z, w), \qquad g(z, w, s)$$

(the first of which is independent of s) are regular on (2) and satisfy

$$(3_n) \qquad |f^n| < L, \ |g| < M \text{ on } (2).$$

Hence, if the Lemma is applied (5_n)-(6_n) (for the fixed n), it follows that the function (4_n) is regular on (7) and satisfies (8_n). This completes the induction.

Next, there exists, in the domain (7), about the origin of the (z, w)-space *a sufficiently small domain,* say

$$|z| < z_0, \qquad |w| < w_0, \tag{11}$$

on which the sequence of functions

$$s_1(z, w), \cdots, s_n(z, w), \cdots \tag{12}$$

is convergent. In the real field, this was proved in a paper of Hartman and myself [2], pp. 862-863, by a refinement of the estimates used by Perron [10] in his case, the case of the Lemma (cf. comments above). In the complex field, the proof is exactly the same as in the real field and will therefore be omitted (just as the proof of the Lemma could be omitted above). But as italicized before (11), both radii z_0, w_0 must be chosen sufficiently small before the estimates of [2] assure the convergence of (12) on (11). In fact, the values of the two positive constants z_0, w_0 had to be subjected in [2], p. 862, to quite a number of inequalities (involving the five positive constants occurring in (2)-(3) above).

This could not be avoided in [2] and, correspondingly, it is not avoided in the present case of the complex field. But the "smallness" of the convergence domain (11) of (12) can, in the present case, be disposed of by an appeal to the oldest theorem (Stieltjes) in the theory of normal families (of functions of two variables, to be sure). In fact, since the functions (4_n) are regular on (7) and, according to (8_n), are uniformly bounded on (7), the convergence of the sequence (12) on a "small" dicylinder (11) implies the convergence of (12) on the entire domain (7), and also the uniformity of the convergence on every (z, w)-domain, say D, the closure of which is contained in (7).

Let (4) denote the limit function of (12) on (7). Then $s(z, w)$ is regular on (7). That the limit function is a solution of (5) (even on a sufficiently small (z, w)-domain), is quite an issue in the real field (cf. the end of the proof in [2], pp. 863-864), since the uniform convergence of the derivatives of the functions (12) must also be assured. There is no such issue in the present case, since the uniform convergence of (12) to $s(z, w)$ on every fixed domain D, defined above, implies that the sequences

$$s_{1z}(z, w), \cdots, s_{nz}(z, w), \cdots \text{ and } s_{1w}(z, w), \cdots, s_{nz}(z, w), \cdots \tag{13}$$

tend to the respective derivatives, $s_z(z,w)$ and $s_w(z,w)$, of $s(z,w)$, uniformly on every fixed domain D. In particular, the three sequences (12), (13) tend to the three functions $s(z,w)$, $s_z(z,w)$, $s_w(z,w)$ at every point (z,w) of the domain (7). Since (4_n) is a solution of (5_n) on (7), it follows that (4) is a solution of (5) on (7). Finally, (6) follows from (6_n), and (8) from (8_n). This completes the proof of the italicized part of Theorem 1.

There remains to be ascertained the parenthetical assertion made after (7). To this end, choose $g(z,w,s)$ to be a function of s alone, and let $f(z,w,s) \equiv 0$. Then (2), (3), (4) and (5), (6) reduce to

$$|s| < c, \quad |g(s)| < M, \quad s = s(z), \quad ds/dz = f(s), \quad s(0) = 0$$

respectively, and (7) can be simplified to the circle $|z| < b/M$. But as shown in [12], there does not exist any *absolute* constant $\epsilon > 0$ in terms of which the circle $|z| < b/M$ could be improved to $|z| < (1+\epsilon)b/M$. For the case in which (5) does not reduce to an ordinary differential equation, cf. the remark to be made in Section 6 concerning the propagation of singularities (characteristics).

6. In contrast to the method of [1], that of [2] consist of successive approximations.* It is the latter method which was used above in order to obtain "sharp" domains of existence in the analytic case, by adapting the considerations of [14].

The propagation of singularities along characteristics, and the nature of the inequalities of Haar (cf. e.g., [4], p. 34 and pp. 119-120) in which this propagation finds it explicit formulation, explain the role played in Theorem 1 by the second of the inequalities (2), the constant L in (2) and (7) being the same as in (3). Correspondingly, it is in the very nature of the problem that something must be lost if the (z,w)-domains considered in assumption (2) and in assertion (7) of Theorem 1 are replaced by (z,w)-dicylinders. The resulting weaker form of Theorem 1 (a weakened

* Since certain recent publications (of none of the authors concerned) make misleading or erroneous statements on a "rediscovery" of the results of Douglis [1] in the last chapter of our paper [2] (a paper the rest of which contains other results also), it should be mentioned here that [1] and [2] were written independently, that the date of receipt of the manuscript of [2] (see p. 834) was December 21, 1951 and that [2] appeared in the October, 1952, issue of the *American Journal of Mathematics*, whereas [1] seems to have appeared sometimes during the summer of 1952 (the exact date and the date of receipt of the manuscript of [1] are not available); so that neither way is it possible to speak of any "rediscovery."

form which cannot, however, be strengthened in terms of (z,w)-dicylinders) is as follows:

THEOREM 2. *Let $f(z,w,s)$, $g(z,w,s)$ be a pair of functions which are regular and bounded, say*

$$|f|<L, \qquad |g|<M, \tag{14}$$

on a tricylinder, say

$$|z|<a, \qquad |w|<b, \qquad |s|<c, \tag{15}$$

about the point $(0,0,0)$ of the space of the complex variables z,w,s. Starting with five such positive numbers L, M, a, b, c, retain the values of L, M and b but replace (if necessary) the given a by a smaller a, and c by a smaller c, so as to satisfy the inequalities

$$0<a<b/L, \qquad 0<c<LM \tag{16}$$

(needless to say, (14) remains true on (15) if the values of a and c, given in (15), are diminished). *Then the solution $s=s(z,w)$ of the quasi-linear initial problem*

$$s_z=f(z,w,s)s_w+g(z,w,s), \qquad s(0,w)\equiv 0 \tag{17}$$

(Cauchy-Kowalewsky) is regular on the dicylinder

$$|z|<\min(a,c/M), \qquad |w|<b-\min(a,c/M) \tag{18}$$

(at least), and

$$|s(z,w)|<c \text{ on } (18). \tag{19}$$

The same is true if $f(z,w,s)$ and $g(z,w,s)$ instead of being regular and subject to (14) on the tricylinder (15), are regular and subject to (14) only on the tricylinder

$$|z|<a, \qquad |w|<b-La, \qquad |s|<c. \tag{20}$$

This is clear from Theorem 1. For, on the one hand, the assumptions (16) assure that (20) and (18) are domains and, on the other hand, the domain (2) is a subset of the tricylinder (20) and the dicylinder (18) is a subset of the domain (7).

7. Instead of the particular case of quasi-linearity, consider now any Cauchy-Kowalewsky equation

$$u_z = F(z, w, u, u_w) \tag{21}$$

with the initial condition

$$u(0, w) \equiv 0 \tag{22}$$

for the solution $u = u(z, w)$ and with an $F = F(z, w, u, v)$ which is given as a regular function of its four variables on a neighborhood of the origin of the complex (z, w, u, v)-space, say on the domain

$$|z| < a, \qquad |w| < b, \qquad |u| < c, \qquad |v| < d. \tag{23}$$

If a, b, c, d are replaced (when necessary) by somewhat smaller positive numbers a, b, c, d, then F is bounded on (23), say

$$|F(z, w, u, v)| < C \text{ on } (23). \tag{24}$$

It follows from (24), and from Cauchy's inequalities for the coefficients of the power series of F, that (21) and (22) are majorized by (21*) and (22) if (21*) denotes the equations which results from (21) if F is replaced by F^*, where

$$F^*(z, w, u, v) = C/\{(1 - z/a)(1 - w/b)(1 - u/c)(1 - v/d)\} \tag{25}$$

on (23). But consider first (21) itself and apply to it a formal device (used already in Kowalewsky's thesis [5]), in a way which reduces (21)-(22) to the quasi-linear case (but in a less drastic form as in [5]), as follows:

First, if $v = v(z, w)$ denotes the partial derivative $u_w = u_w(z, w)$, then

$$u(0, w) \equiv 0, \qquad v(0, w) \equiv 0, \tag{26}$$

by (22). and partial differentiation of (21) with respect to w shows that

$$\begin{aligned} u_z &= F(z, w, u, v) + 0 . u_w, \\ v_z &= H(z, w, u, v) + F_v(z, w, u, v) . v_w, \end{aligned} \tag{27}$$

where

$$H(z, w, u, v) = F_w(z, w, u, v) + vF_u(z, w, u, v), \tag{28}$$

since $u_{zw} = u_{wz}$. But the comparison of like powers of z, w shows that there is a *unique* pair of (formal) power series u, v in (z, w) satisfying (27) and (26), and that there is a *unique* (formal) power series in (z, w) satisfying (21) and (22). Consequently, (27) and (26) together are not only necessary but sufficiently as well for (21) and (22) together.

This implies that if (28*) and (27*) denote the relations which result if F, H are replaced by F^*, H^* in (28) and (27), respectively, then (21*) and (22) together are equivalent to (27*) and (26) together. Finally, it is clear from (27*) that if the (non-negative) numbers γ_{hijk} are defined by placing, for any fixed set of indices, $\gamma_{hijk} = \max(\alpha_{hijk}, \beta_{hijk})$, where

$$F^* = \Sigma\, \alpha_{hijk} z^h w^i u^j v^k, \qquad H^* = \Sigma\, \beta_{hijk} z^h w^i u^j v^k$$

on (23) (with $\Sigma = \sum_{h=0}^{\infty} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty}$), and if G is then defined on (23) by placing

$$G(z, w, u, v) = \Sigma\, \gamma_{hijk} z^h w^i u^j v^k, \tag{29}$$

then the case

$$f(z, w, s) = F^*_v(z, w, s, s), \qquad g(z, w, s) = G(z, w, s, s) \tag{30}$$

of the scalar conditions (17) is a majorant of the components of the vector conditions (27*), (26).

It follows that if Theorem 2 is applied to the case (30) of (17), then there results a pair of positive numbers, say A and B, which depend only on the five constants occurring in (23) and (24) and which have the property that the solution $u = u(z, w)$ of (21) and (22) must be regular on the dicylinder

$$|z| < A, \qquad |w| < B. \tag{31}$$

Since (31) represent (18), it would be easy to write down the explicit form, in terms of a, b, c, d and C, of such a pair of constants A, B. This will however be omitted, since the resulting dicylinder (31) cannot possibly supply a "best" dicylinder. Incidentally, since the method involes majorizations, it leads to explicit estimates of the domain of regularity of the solutions apply also in the case of Cauchy-Kowalewsky *systems* (21)-(22), where u, F and w are vectors (but z is not).

8. The following fact is mentioned because of its methodical interest.

THEOREM 3. *If a function $F(z, w, u, v)$ is regular on a neighborhood (23) of the origin of the complex (z, w, u, v)-space, then there exists a pair of positive constants, say α and β, having the following properties: The functions*

$$u^1(z, w), \cdots, u^n(z, w), \cdots, \tag{32}$$

which are assigned by $u^0(z, w) \equiv 0$ and the "successive approximations" algorithm

$$u_z^n(z,w) = F(z,w,u^n(z,w),u_w^{n-1}(z,w)), \qquad u^n(0,w) \equiv 0, \tag{33}$$

exist and are regular on the dicylinder

$$|z| < a, \qquad |w| < \beta \qquad (a \leqq a, \beta \leqq b) \tag{34}$$

(which is independent of n), the inequalities

$$|u^n(z,w)| < c \text{ and } |u_w^n(z,w)| < d \text{ on } (34) \tag{35}$$

hold for every n and for the constants c, d occurring in (23), *and the sequence* (32) *converges on* (34) *to the solution $u = u(z,w)$ of the Cauchy-Kowalewsky problem* (21)-(22).

The truth of the assertions of Theorem 3 becomes clear if Section 7 is compared with the proof given in Section 5.

Theorem 3 supplies for the Cauchy-Kowalewsky theorem of a single differential equation a proof which goes much deeper than either of the classical proofs (those based either on the *calcul des limites* or on the use of characteristics, as in the real domain). The possibility of proving the theorem via Theorem 3 is more than a mere curiosity, since the method of successive approximations usually supplies an existence range going far beyond the existence range supplied by the other methods. In this regard, cf. the comments of Painlevé [7] and the results of [13].

9. It is worth mentioning that, from the formal point of view of "integrability," the varying degrees of generality, represented by the cases of Theorem 3, Theorem 1 and the Lemma of Section 4, lead to "reductions" of the system of ordinary differential equations to which the method of characteristics, referred to in Section 8, is based.

First, the characteristic equations belonging to (21) are

$$dw/dz = -F_v, \qquad du/dz = F - vF_v, \qquad dv/dz = F_w + vF_u, \tag{36}$$

where $F = F(z,w,u,v)$. Next, if (21) is quasi-linear, as in (5), then $F = vf + g$, where are f and g functions of (z, w, u) only, then the ternary system (36) *splits* into the binary system

$$dw/dz = f(z,w,u), \qquad du/dz = g(z,w,u) \tag{37}$$

and into the equation

$$dv/dz = \alpha(z,w,u) + \beta(z,w,u)v + \gamma(z,w,u)v^2, \tag{38}$$

where $\alpha = g_w$, $\beta = g_u - f_z$, $\gamma = -f_u$; so that, if a solution $(w, u) = (w(z), u(z))$ of (37) is known, then (38) reduces to a *Riccati equation*

(38 bis) $$dv/dz = a(z) + b(z)v + c(z)v^2$$

for $v = v(z)$. Finally, if $f(z, w, u)$ is assumed to have Perron's particular form, that is, if $f = f(z, w)$ (cf. the Lemma of Section 4), then the binary system (37) *splits* into two equations of first order:

(37_1) $$dw/dz = f(z, w), \qquad (37_2) \qquad du/dz = g(z, w, u).$$

In fact, if a solution $w = w(z)$ of (37_1) is known, then (37_2) becomes of the form

(37_2 bis) $$du/dz = h(z, u),$$

and is therefore, like (37_1), a differential equation for a single function.

10. The following considerations deal, like Section 7, with the Cauchy-Kowalewsky majorant (21*)-(22) of (21)-(22); cf. (25) and (24). But the method will be more direct, and the result more explicit, than in Section 7.

Let the notation be so chosen that all four radii occurring in (23) become 1. Then (24) and (25) show that the Cauchy-Kowalewsky majorant (21*) of (21) becomes

(40) $$\phi_z = C/\{(1-z)(1-w)(1-\phi)(1-\phi_w)\},$$

where C is a positive constant. Let $\phi(z, w)$ denote that solution of (40) which is determined by the initial condition

(41) $$\phi(0, w) \equiv 0;$$

cf. (21) and (22), where $u = \phi$.

According to Perron [9], p. 158, the function $\phi(z, w) = \phi_C(z, w)$ is regular in the dicylinder

(42) $$|z| < (1-r)^2/(1+8C), \qquad |w| < r,$$

at least, if r is any number on the interval $0 < r < 1$. It is however natural to expect that, in view of (40), the location of the singularities will be such that *every point* (z, w) *of the dicylinder* $(|z| < 1, |w| < 1)$ *becomes a regular point of* $\phi_C(z, w)$ *as* $C \to 0$. But (42) it too rough to prove this. It will be shown, however, that (42) can be improved to

(43) $$|z| < 1 - \exp(-\tfrac{1}{8}(1-r)^2/C), \qquad |w| < r.$$

This result (the proof of which will depend on an application of (47) below) proves the italicized desideratum, since if $r=1-\epsilon<1$ is fixed and $C=C_\epsilon$ is large enough, then the dicylinder (43) contains the dicylinder ($|z|<1-\epsilon$, $|w|<1-\epsilon$). It will be seen in what follows that the *structure of* (43) *is "correct,"* in the sense that (43) is close to the ultimate truth.

In order to formulate this situation precisely, let C be fixed in (40) and let r be any positive number corresponding to which there exists a positive R having the property that the solution $\phi(z,w)$ of (40) and (41) is regular in the dicylinder ($|z|<R, |w|<r$). By $R=R(r)$ will be meant the greatest such R (so that $(R(r),r)$ actually is a pair of associated radii, a pair to which Hartogs's convexity theorem applies).

THEOREM 4. *If $R(r)=R_C(r)$ denotes the z-radius associated to a w-radius $r>0$ in the solution $\phi=\phi(z,w)=\phi_C(z,w)$ of* (40) *and* (41), *then, on the one hand*

$$R(r)=0 \text{ if } r=1 \text{ (hence, if } r>1) \tag{44}$$

and, on the other hand,

$$R(r)=1-\exp(-(1-r)S(r)/C) \text{ if } 0<r<1, \tag{45}$$

where $S(r)$ is a function satisfying the inequalities

$$\tfrac{1}{8}(1-r)\leqq S(r)\leqq \tfrac{1}{2} \qquad (0<r<1) \tag{46}$$

(*cf. also* (49) *below, where* $\gamma>0$).

Both of the equality signs can be excluded in (46), since it will be clear from the proof of (46) that neither of the constants $\frac{1}{2}$, $\frac{1}{8}$ can be "best"; cf. also (49) below. Note that C occurs explicitly in (45) but not in (46), even though it is not clear that $S(r)$ is not a function, $S_C(r)$, of C also. Since (45) and the second of the inequalities (46) imply that $R(r)\to 0$ as $r\to 1$, they imply (44).

11. First, since the coefficients of all the power series involved are positive (or 0), it is clear that a minorant of (40) and (41) results if (40) is retained and (40) is replaced by the partial differential equation, say (40 bis), which results if the fourth factor, $(1-\phi_w)$, of the denominator is omitted in (40). But (40 bis) can be written in the form

$$\phi_z=\mu/\{(1-z)(1-\phi)\}, \qquad \mu=C/(1-w), \tag{40 bis}$$

where μ is a parameter, and so (40 bis) is an (ordinary, rather than a partial) differential equation which the substitution

$$t = -\log(1-z) \qquad (|z|<1,\ \log 1 = 0) \tag{47}$$

transforms into $d\phi/dt = \mu(1-\phi)$. The solution $\phi = \phi(t)$ is given by $(1-\phi)^2 = -2\mu t + \text{const.}$, where const. $= 0$ by virtue of (41) and (47). It follows therefore from (47) that the solution $\phi = \phi(z;w)$ of (40 bis) becomes singular at the z-value $z_0 = z_0(\mu)$ determined by the equation $(1-z_0)^{2\mu} = 1/e$. Since $\mu = C/(1-w)$, this equation leads to

$$z_0 = 1 - \exp(-\tfrac{1}{2}(1-r)/C) \tag{48}$$

if $w = r$, where $0 < r < 1$.

Since (40) is a majorant of (40 bis) (with reference to the common initial condition (41)), it follows that the solution $\phi(z,w)$ of (40) and (41) cannot be regular on the closure of the dicylinder $(|z| < z_0, |w| < r)$ if z_0 is the value determined by (48), where $0 < r < 1$, hence $0 < z_0 < 1$. In view of the definition of $R(r)$, this proves the second of the inequalities (46) for the function $S(r)$ defined by (45). Incidentally, it is easy to see that the $\frac{1}{2}$ in (46) can be improved to $\frac{1}{3}$ if the factor $1/(1-\phi_w)$ in (41) is relaxed only to $1/(1-\phi)$ and not, as in the replacement of (40) by (40 bis), to $1/(1-0) = 1$. But this is unimportant, since what is missing in (46) is the improvement of the constant upper bound for $S(r)$ to an upper bound which depends on r (cf. Section 11 below), in the same way as the lower bound supplied by (46). In view of (45), this lower bound for $S(r)$ is equivalent to the result (43) of [13].

Concerning an r-dependent improvement of the upper bound in (46), I find, by using minorants of (40) which are more elaborate than (40 bis), that

$$S(r) = O(1-r)^{\gamma} \text{ as } r \to 1 \tag{49}$$

holds for some constant $\gamma > 0$ (for instance, for $\gamma = \frac{2}{3}$) and the computations seem to indicate that (49) holds for any $\gamma < 1$. If (49) should be true for $\gamma = 1$ also, then, writing (45) in the form

$$R(r) = 1 - \exp(-(1-r)^2 T(r)/C), \text{ where } 0 < r < 1, \tag{50}$$

one could conclude from the lower bound supplied by (46) that

$$0 < \operatorname*{fin\,inf}_{0<r<1} T(r) \leqq \operatorname*{fin\,sup}_{0<r<1} T(r) < \infty\ ; \tag{50?}$$

so that the problem of the pair of the associated radii $(R(r), r)$ of $\phi(z,w)$

would become refined to the determination of the two positive values occurring in (50?).

12. There remains to be proved the assertion of Theorem 4 concerning the first of the inequalities (46). In view of (45), that assertion can be formulated as follows: If r is any number on the interval $0 < r < 1$, then the solution $\phi = \phi(z, w)$ of (40) and (41) is regular on the dicylinder

$$|z| < 1 - \exp(-(1-r)^2/8C), \qquad |w| < r \tag{51}$$

(and, therefore, on the union of the r-family (51) of dicylinders).

Perron [9] obtained his domain (42) by (retaining (41) but) *first* majorizing (40) by

$$\phi_z = C/\{(1-z)(1-w)(1-(1-w)^{-1}\phi - \phi_w)\} \tag{52}$$

and *then* (52) by

$$\phi_z = C/\{(1-z)(1-[1-z]^{-2}w)(1-[1-w]^{-1}\phi - \phi_w)\}. \tag{53}$$

But it will be seen that the device, applied by Perron to an "explicit" integration of (53), can be adjusted so as to work for (52) itself, rather than just for the weakened form, (53), of (52). This will be made possible by an application to (52) of the mapping (47). The result will be that the solution $\phi(z, w)$ of (52) and (41) is regular on any dicylinder (51), where $0 < r < 1$.

First, (47) transforms (52) into

$$(1-w)\phi_t = C/\{1-(1-w)^{-1}\phi - \phi_w\} \tag{54}$$

but leaves (41) unaltered, since $z = 0$ corresponds to $t = 0$ in (47). Next, replace $\phi = \phi(t, w)$ by $\psi = \psi(t, w)$, where

$$\psi = \phi/(1-w). \tag{55}$$

Then it turns out that, just as in Perron's case (cf. [9], pp. 157-158), the function $\psi = \psi(t, w)$ (determined, near $(t, w) = (0, 0)$, by (53), (54) and (41)), is a function $\psi = \psi(u)$ of a *single* variable u, to be defined by

$$u = t/(1-w)^2. \tag{56}$$

In fact, a straightforward calculation shows that, by virtue of (55) and (56), the *partial* differential equation (54) is identical with the *ordinary* differential equation

$$\psi' = C/(1-2u\psi'), \tag{57}$$

where $\psi = \psi(u)$ and $\psi' = d\psi/du$, whilst (41) goes over into the initial condition $\psi(0) = 0$.

It is clear that (57) is a quadratic equation for $\psi' = \psi'(u)$, and that its root ψ' which remains regular at $u = 0$ is

$$\psi' = (1 - (1 - 8Cu)^{\frac{1}{2}})/4u. \tag{58}$$

The function (58) of u is regular on the circle

$$|u| < (8C)^{-1} \tag{59}$$

of the u-plane. Since $\psi = \psi(u)$ follows from (58) (and from the initial condition $\psi(0) = 0$) by a quadrature, $\psi(u)$ itself must be regular on the circle (59). This completes the proof. For, on the one hand, $\phi(z, w)$ is identical with $\psi(u)$ by virtue of (47), (55) and (56), and, on the other hand, it is readily seen from (47) and (56) that the u-circle (59) corresponds to a (z, w)-domain which contains the dicylinder (51) belonging to an arbitrary value of r, where $0 < r < 1$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Douglis, "Existence theorems for hyperbolic systems," *Communications of Pure and Applied Mathematics*, vol. 5 (1952), pp. 119-159.

[2] P. Hartman and A. Wintner, "On hyperbolic partial differential equations," *American Journal of Mathematics*, vol. 74 (1952), pp. 834-864.

[3] E. Kamke, *Differentialgleichungen reeller Funktionen*, Leipzig, 1930.

[4] ———, *Differentialgleichungen*, Lösungmethoden und Lösungen, vol. II, Leipzig, 1944.

[5] S. von Kowalewsky, "Zur Theorie der partiellen Differentialgleichungen," *Crelle's Journal*, vol. 80 (1875), pp. 1-32.

[6] P. D. Lax, "Nonlinear hyperbolic equations," *Communications of Pure and Applied Mathematics*, vol. 6 (1953), pp. 231-258.

[7] P. Painlevé, "Sur le calcul des intégrales d'un système différentiel par la méthode de Cauchy-Lipschitz," *Bulletin de la Société Mathématique de France*, vol. 27 (1899), pp. 149-152.

[8] J. Perausówna, "Sur le domaine d'existence des intégrales de l'équation $p + f(x, y, z)q = g(x, y, z)$," *Annales de la Societé Polonais de Mathématique*, vol. 12 (1933), pp. 1-5.

[9] O. Perron, "Ein neuer Bewis des Fundamentalsatzes in der Theorie der partiellen Differentialgleichungen erster Ordnung," *Mathematische Zeitschrift*, vol. 5 (1919), pp. 154-160.

[10] ———, "Ueber Existenz und Nichtexistenz von Integralen partieller Differentialgleichungenssyteme im reellen Gebiet," *ibid.*, vol. 27 (1928), pp. 549-564.

[11] T. Ważewski, "Sur le domaine d'existence des intégrales de l'équations aux derivées partielles du premier ordre linéaire," *Annuales de la Societé Polonaise de Mathématique*, vol. 12 (1933), pp. 6-15.

[12] A. Wintner, "On the exact value of the bound for the regularity of solutions of ordinary differential equations," *American Journal of Mathematics*, vol. 57 (1935), pp. 539-540.

[13] ———, "On the method of successive approximations in initial value problems," to appear in the *Annali di Matematica*, 1956.

ON CERTAIN ABSOLUTE CONSTANTS CONCERNING ANALYTIC DIFFERENTIAL EQUATIONS.*

By Aurel Wintner.

PART I.

This note on *ordinary* differential equations, though it does not presuppose the paper on *partial* differential equations which precedes it in this volume (pp. 525-541), can be thought of as an appendix to that paper, since it contains (among other things) a justification of the choice of the region of analyticity in Theorem 1 *loc. cit.*

Consider the solution $w = w(z)$ of the differential equation

$$dw/dz = f(z, w) \tag{1}$$

with the initial condition

$$w(0) = 0 \tag{2}$$

and under the following assumptions: $f(z, w)$ is regular on the dicylinder

$$|z| < a, \qquad |w| < b \tag{3}$$

and is bounded there, say

$$|f(z, w)| < M \text{ on } (3). \tag{4}$$

Then a classical result states that the method of successive approximations supplies the circle

$$|z| < \min(a, b/M) \tag{5}$$

on which the solution $w(z)$ of (1) and (2) is sure to be regular.

For a long time, and by methods usually distinct from that of the successive approximations, various efforts have been made in order to improve on the radius of the circle (5) of assured regularity for $w(z)$; cf. M. Müller's report in [6], pp. 169-172. But it was shown in [8] that those efforts

---

* Received February 29, 1956.

could not possibly succeed, since, if $f(z, w)$ is independent of z (so that (1) and (5) reduce to

$$dw/dz = f(w) \tag{6}$$

and $|z| < b/M$ respectively), then, corresponding to every $\epsilon > 0$, it is possible to exhibit an $f(w) = f_\epsilon(w)$ having the property that the solution $w(z)$ of (6) and (2) will possess a singularity within the circle $|z| < (1+\epsilon)b/M$.

In this sense, (5) is the best possible result. But as mentioned in [8], this result depends on the assumption that no supplementary information, such as a Lipschitz constant, is added to the estimate (4). The purpose of the following considerations is to fill in the resulting gap.

Let $f(z, w)$ be regular on (3) but, instead of assuming (4), suppose only that $f(z, 0)$ is bounded, say

$$|f(z, 0)| \leqq N \text{ on } |z| < a, \tag{7}$$

but suppose also that the partial derivative of $f(z, w)$ with respect to w is bounded, say

$$|\partial f(z, w)/\partial w| < L \text{ on } (3). \tag{8}$$

Although the field is complex, it is readily seen that (8) is equivalent to Lipschitz's condition

$$|f(z, w') - f(z, w'')| < L\,|w' - w''|, \tag{8 bis}$$

where (z, w'), (z, w'') are any two points (z, w) on (3) with distinct w but common z. Since (8) is equivalent to (8 bis), it follows from a remark of Painlevé [7] on a result of E. Lindelöf [4], p. 123 (concerning successive approximations), that the solution $w(z)$ of (1) and (2) is regular on the circle

$$|z| < \min\,(a,\ L^{-1}\log(1 + bL/N)) \tag{9}$$

(at least). Before Lindelöf (but not with the method of successive approximation), and in the real field, the domain (9) was found by Lipschitz himself ([5], pp. 509-514). Cf. also a paper of O. Hölder [2].

The choice of the radius of the circle (9), in contrast to that of (5), seems to be artificial (except, perhaps, if an appeal is made to the inequality of Haar [1] in the theory of characteristics). But it turns out that (9), like (5), is the best possible result of its own kind. By this is meant that, if either the a or the $L^{-1}\log$ in (9) is multiplied by $1+\epsilon$, where $\epsilon > 0$ is arbitrarily small, then there results a z-circle within which the solution $w(z)$ of (1) and (2) will acquire a singularity, if $f(z, w) = f_\epsilon(z, w)$ is suitably chosen.

This is clear for the replacement of a by $(1+\epsilon)a$ in (9). In fact, if $f(z,w)$ is independent of w, then (1) and (9) reduce to $dw/dz = f(z)$ and $|z| < a$ respectively, and so the solution $w(z)$ of (1) and (2) will become singular on the boundary $|z| = a$ whenever $f(z)$ does. Consequently, it is sufficient to consider the case complementary to the case $f(z,w) = f(z)$, that is, the case $f(z,w) = f(w)$.

Then, if $b = 1$ without loss of generality, (1) and (9) reduce to (6) and

$$|z| < L^{-1}\log(1 + L/|f(0)|) \tag{10}$$

respectively, since the N in (7) can be chosen to be the value of $|f(w)|$ at $w = 0$ (this value can be assumed to be distinct from 0, since $w(z) \equiv 0$ is the solution of (6) and (2) if $f(0) = 0$). On the other hand, the formulation (8) of Lipschitz's condition (8 bis) reduces to the assumption that

$$|f'(w)| < L \text{ for } |w| < 1, \tag{11}$$

where $f' = df/dw$ and $1 = b$. Accordingly the assertion, to be proved, is as follows: If it is only assumed that the function $f(w)$ is regular on the circle $|w| < 1$ and that its derivative is restricted by (11), then the solution $w(z)$ of (6) and (2) *can* become singular at $z = z^0$ whenever $|z^0|$ is given as a number exceeding the radius of the circle (10).

Choose $f(0) = 1$ and write ϵ instead of L. Then $f(w)$ is a power series of the form

$$f(w) = 1 + g(w), \text{ where } g(w) = \sum_{n=1}^{\infty} c_n w^n \tag{12}$$

for $|w| < 1$, the circle (10) reduces to

$$|z| < \epsilon^{-1}\log(1+\epsilon) = 1 + o(1) \text{ if } \epsilon \to 0, \tag{13}$$

and condition (11) is satisfied (for a fixed ϵ) if

$$\sum_{n=1}^{\infty} n\,|c_n| < \epsilon, \tag{14}$$

where $c_n = c_n(\epsilon)$. On the other hand, it is clear from (12) that, if ϵ is fixed in $g(w) = g_\epsilon(w)$, then, near $z = 0$, the solution $w = w(z) = w_\epsilon(z)$ of (6) and (2) is the (local) inverse of the function $z = z(w) = z_\epsilon(w)$ defined by

$$z(w) = \int_0^w (1 + g(w))^{-1}\,dw, \text{ where } g(0) = 0. \tag{15}$$

The only restriction on the coefficients $c_n = c_n(\epsilon)$ of the power series (12) is that they should be "small" in the sense of (14). Hence it is easily realized (cf. a parallel construction in the proof of assertion (β^*) of (ii) below) that, for every $\epsilon > 0$, it is possible to choose the coefficients $c_n = c_n(\epsilon)$ of $g(w) = g_\epsilon(w)$ in accordance with (14) and in such a way that the function $w(z) = w_\epsilon(z)$, defined near $z = 0$ as the inverse of the function (15), becomes singular at some point of a circle $|z| < r_\epsilon$ the radius of which is of the form $r_\epsilon = 1 + o(1)$ as $\epsilon \to 0$. But this is precisely the assertion, since the circle (10) is represented by (13).

PART II.

If $a = 1$, $b = 1$ and $M = 1$ in (3)-(4), then the circle (5) becomes $|z| < 1$ and, in view of the maximum principle. nothing is lost if the sign of equality is allowed in (4). If this is combined with the circumstance that successive approximations $w_0(z) \equiv 1$, $w_1(z), \cdots, w_n(z), \cdots$ which lead to the solution $w(z)$ of (1) and (2) are subject to the inequality $|w_n(z)| < b$ on the circle (5), then there results the following fact (i):

(i) *If $f(z, w)$ is regular, and satisfies the inequality $|f(z, w)| \leqq 1$, on the dicylinder $(|z| < 1, |w| < 1)$, then the solution $w(z)$ of* (1) *and* (2) *is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < 1$.*

For a refinement of (i), cf. [14].

Other applications of known "best contants" pertaining to classes of power series result if (6) is contrasted with

$$dw/dz = f^*(w), \tag{16}$$

where

$$f^*(w) = \sum_{n=0}^{\infty} |c_n| w^n \text{ if } f(w) = \sum_{n=0}^{\infty} c_n w^n. \tag{17}$$

In order to see this, suppose that $f(w)$, hence $f^*(w)$, is regular on the circle $|w| < 1$. Then it is clear from the proof of (i) that the solution $w(z)$ of (16) and (2) is regular on any circle

$$|z| < r/f^*(r), \tag{18}$$

where r is any positive number less than 1 (for a direct proof, cf. Lindelöf's method in [11]).

Suppose that $f(w)$, besides being regular for $|w| < 1$, satisfies the inequality

$$|f(w)| < 1 \text{ for } |w| < 1. \tag{19}$$

Then, according to Bohr (cf. [10], pp. 32-34), the inequality $f^*(r) \leqq 1$ holds if $r = \frac{1}{3}$ but not in general if $r = \frac{1}{3} + \epsilon > \frac{1}{3}$. It follows from the first of these two facts (the second fact, that concerning the non-existence of an absolute constant $\epsilon > 0$, is irrelevant this time) that the radius of the circle (18) becomes not less than $\frac{1}{3}$ at $r = \frac{1}{3}$. Consequently, (19) is sufficient in order that the solution $w(z)$ of (16) and (2) be regular on the circle $|z| < \frac{1}{3}$.

On the other hand, as will be shown elsewhere [16], even the most favorable choice of r (on the range $0 < r < 1$) can bring the radius of (18) arbitrarily close to $\frac{1}{3}$ if only (19) is assumed. One might therefore expect that $\gamma = \frac{1}{3}$ is the greatest absolute constant having the property that the solution $w(z)$ of (16) and (2) is regular on the circle $|z| < \gamma$ whenever (19) holds. This expectation proves, however, to be erroneous. The explanation is that $\frac{1}{3}$ is the best constant by virtue of Cauchy's principle of "majorization," but *not* by virtue of the finer principle of "subordination" (in this regard, cf. the concluding remarks in [15]). In fact, it turns out that $|z| < \frac{1}{3}$ can be improved to $|z| < \gamma$ (but not, of course, to $|z| < 1$), where γ denotes the least positive number r (< 1) for which the function

$$r(1-r^2)^{-\frac{1}{2}} - \arcsin r \qquad (0 \leqq r < 1) \tag{21}$$

becomes 1 (in (21), both the square root and the arc sin are meant to be positive).

(I) *If $f(w)$ is regular on the circle $|w| < 1$ and satisfies* (19), *then, while the solution $w(z)$ of* (6) *and* (2) *is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < 1$, the solution $w(z)$ of* (16) *and* (2) *is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < \gamma$, where γ denotes the least positive root of the transcendental equation*

$$\gamma/(1-\gamma^2)^{\frac{1}{2}} = 1 + \arcsin \gamma \qquad (\arcsin 0 = 0). \tag{22}$$

The first of the two assertions of this theorem (I) is contained in (i). In order to prove the second assertion of (I), note that, according to (17),

$$|f^*(w)|^2 \leqq \sum_{n=0}^{\infty} |c_n|^2 \sum_{n=0}^{\infty} |w|^{2n}, \text{ where } \sum_{n=0}^{\infty} |c_n|^2 < 1,$$

by (19), except when $f(z) = cz$, where $|c| = 1$. Hence, in any case,

$$f^*(r) < (\sum_{n=0}^{\infty} r^{2n})^{\frac{1}{2}} = (1-r^2)^{-\frac{1}{2}} \text{ if } 0 < r < 1, \tag{23}$$

where $f^*(r) \geqq 0$, by (17).

Recourse can now be had to a theorem on "subordination," initiated by Nakano [12] and formulated in a general form in [15]. In fact, it is clear from (23) and from the lemma of [15], p. 1106, that the solution $w(z)$ of (16) and (2) is regular, and satisfies the inequality $|w(z)|<1$, on the circle $|z|<\gamma$, if γ is any positive number having the following property: If the function $r=r(t)$ $(\geqq 0)$ is defined as the solution of the differential equation

$$dr/dt=(1-r^2)^{-\frac{1}{2}} \tag{24}$$

and of the initial condition

$$r(0)=0, \tag{25}$$

then $r(t)$ exists, and satisfies the inequality $r(t)<1$, on the interval $0\leqq t<\gamma$. But (24) can be solved by the inversion of a quadrature, whence it is seen that the solution $r=r(t)$ of (24) and (25) is the (local) inverse of the function $t(r)$ which results if the function (21) of r is denoted by $t=t(r)$. Finally, it is clear that the value of the function (21) is between 0 and 1 for $0<r<\gamma$, if γ is the least positive root (<1) of the equation (22). This completes the proof of (I).

Since the estimate (23) of $f^*(r)$ by $(1-r^2)^{-\frac{1}{2}}$ excludes the sign of equality (except, perhaps, at $r=0$), it also follows that *the* γ *of* (I) *can be improved to* $\gamma+\epsilon$, *where* $\epsilon=\epsilon_f>0$. On the other hand, it remains undecided whether this ϵ can be chosen independent of f (in other words, whether γ is the best absolute constant). In this regard, nothing seems to follow from a direct consideration of the sequence of rational functions which, for quite another purpose, Landau has constructed from the sequence of the partial sums of

$$(1-W)^{-\frac{1}{2}}=\sum_{n=0}^{\infty} b_n W^n, \text{ where } b_n=1.3\cdots(2n-1)/2.4\cdots 2n \tag{23 bis}$$

($W=w^2$ in (20); cf. [10], pp. 26-29, and a general theorem of Schur [13], pp. 122-124).

It will now be shown that if the circle $|z|<\gamma$, defined in (I), is replaced by the smaller circle $|z|<\sin\gamma$, then (I) can be transferred to the case in which (16) is generalized to

$$dw/dz=f^*(z,w), \tag{26}$$

where, corresponding to (17),

$$(27) \qquad f^*(z,w) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |c_{mn}| z^m w^n \text{ if } f(z,w) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} c_{mn} z^m w^n$$

(so that $f^*(z,w)$ is regular on a dicylinder (3) whenever $f(z,w)$ is).

(II) *If $f(z,w)$ is regular, and satisfies the inequality $|f(z,w)| < 1$, on the dicylinder ($|z| < 1, |w| < 1$), then the solution $w(z)$ of (26) and (2) is regular, and satisfies the inequality $|w(z)| < 1$, on the circle $|z| < \sin\gamma$, where γ (< 1) denotes the least positive root of the transcendental equation (22).*

In view of (i), the content of (II) is that, if (1) is replaced by (26), then the circle $|z| < 1$, supplied by (i), becomes replaced by a circle $|z| < \lambda$, where λ is an absolute constant (< 1) which (II) claims to be not less than $\sin\gamma$. It can readily be shown that, corresponding to the remarks made before (I) on the $\frac{1}{3}$-radius, the simple method of Lindelöf [11] supplies for (26) and (2) an absolute constant λ which is substantially less favorable than the absolute constant supplied by (II). The proof of (II) proceeds as follows:

Clearly, the assumptions of (II) and the second of the relations (27) imply that, corresponding to the proof of (20),

$$\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |c_{mn}|^2 \leqq 1; \text{ hence } f^*(|z|, |w|) \leqq \left(\sum_{m=0}^{\infty} \sum_{n=0}^{\infty} |z|^{2m} |w|^{2n}\right)^{\frac{1}{2}}$$

if $0 < |z| < 1$, $0 < |w| < 1$, by the first of the relations (27). Since the preceding square root is $(1-s^2)^{-\frac{1}{2}}(1-r^2)^{-\frac{1}{2}}$ if $|z| = s < 1$, $|w| = r < 1$, it follows that (26) and (2) are "subordinated" to

$$(28) \qquad dr/ds = (1-s^2)^{-\frac{1}{2}}(1-r^2)^{-\frac{1}{2}}$$

and (25). In fact, if [15] is applied in the same way as in the proof of (I), it follows that the solution $w(z)$ of (26) and (2) is regular, and in absolute value less than 1, on any circle $|z| < \lambda$ the radius of which is a positive number satisfying $\lambda \leqq 1$ and having the following property: The solution $r = r(s)$ ($\geqq 0$) of (26) and (25) exists, and satisfies the inequality $r(s) < 1$, on the interval $0 \leqq s < \lambda$.

On the other hand, if

$$(29) \qquad t = \arcsin s, \text{ where } 0 \leqq t < \tfrac{1}{2}\pi, \qquad 0 \leqq s < 1,$$

then t is increasing with s (and $s = 0$ corresponds to $t = 0$), and (29) transforms (28) into (24) and leaves (25) unaltered. But the solution $r = r(t)$ of (24) and (25) exists, and satisfies the inequalities $0 \leqq r(t) < 1$, on the

interval $0 \leqq t < \gamma$ (cf. the end of the proof (I)). It follows therefore from (29) that the solution $r = r(s)$ of (28) and (25) exists, and satifies the inequalities $0 \leqq r(s) < 1$, on the interval $0 \leqq s < \lambda$, where $\lambda = \sin\gamma$. In view of the end of the preceding paragraph, this completes the proof of (II).

It is instructive to compare the above proofs and results with those of Cauchy himself (cf., e.g., pp. 6-7 of Bieberbach's book (1953) on the complex function theory of ordinary differential equations). Cauchy assumes that $f(z, w)$ is regular and bounded on a dicylinder (3). It can be assumed that $a = 1$ and $b = 1$ in (3), and that (4) holds for $M = 1$. Then the assumptions on $f(z, w)$ are precisely the same as in (i) or (II). But instead of Parseval's relation and Schwarz's inequality, used in the proof of (II) above, and instead of using the principle of *subordination* (which is the crucial step in the above proof; cf. [15], p. 1107), Cauchy applies his inequalities, $|c_{mn}| \leqq M/a^m b^n$, for the coefficients of (27), which commits him to the following *majorization* of the power series (27):

$$|f(z, w)| \leqq f^*(|z|, |w|) \leqq \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} 1 \cdot |z|^m |w|^n,$$

since $M/a^m b^n = 1$ in the present normalization. Accordingly, not only (1) but also (26) is being majorized by

$$dw/dz = (1-z)^{-1}(1-w)^{-1} \tag{30}$$

and, correspondingly, (28) is roughened to

$$dr/ds = (1-r)^{-1}(1-s)^{-1} \tag{31}$$

(leaving no possibility for applying "subordination" instead of "majorization").

The substitution $Z = -\log(1-z)$, where $|z| < 1$ (and $Z = 0$ at $z = 0$), reduces (30) to $dw/dZ = (1-w)^{-1}$. The solution of the latter differential equation and of (2) is $w = \frac{1}{2} - \frac{1}{2}(1-2Z)^{\frac{1}{2}}$, a function which is regular on the circle $|Z| < \frac{1}{2}$ but not at the point $Z = \frac{1}{2}$. Hence it is clear from $z = 1 - e^{-Z}$ that the solution $w(z)$ of (30) and (2) is regular on the circle

$$|z| < \lambda, \text{ where } \lambda = 1 - e^{-\frac{1}{2}}, \tag{32}$$

and becomes singular at the boundary point $z = \lambda$ of the circle (32). It follows that the solution $w(z)$ of (26) and (2) is regular on the circle (32).

But this circle is substantially smaller than the circle, $|z| < \sin\gamma$, supplied by (II).

It may finally be mentioned that the discrepancy between the respective radii, γ and $\sin\gamma$ $(<\gamma)$, supplied by (I) and (II), has an analogue when only Cauchy's majorants are applied. In fact, if (26) is of the particular form (16), then (30) can be replaced by $dw/dz = (1-w)^{-1}$. Since the solution of the latter differential equation and of (2) is $w = \frac{1}{2} - \frac{1}{2}(1-2z)^{\frac{1}{2}}$, it follows that Cauchy's method of majorants improves his circle (32) to $|z| < \frac{1}{2}$ (though not to any circle $|z| < \frac{1}{2} + \epsilon$, where $\epsilon > 0$), if (26) is of the particular form (16).

In the particular case (6) of (1), the information contained in (i) can be completed as follows:

(ii) *On the circle* $|w| < 1$, *let* $f(w)$ *be a regular function*

$$f(w) = \sum_{n=0}^{\infty} c_n w^n \tag{33}$$

satisfying (19). *Denote by* $w(z) = w_f(z)$ *the solution of* (6) *and* (2). *Then*

(α) *the function* $w(z)$ *is regular not only on the circle* $|z| < 1$ *but also on some circle* $|z| < 1 + \epsilon$, *where* $\epsilon = \epsilon_f > 0$ (*and, except when* $f(w) \equiv f(0)$ *and* $|f(0)| = 1$, *not only* $|w(z)| < 1$ *but also* $|w(z)| < \text{const.} < 1$ *holds for* $|z| < 1$) *but*

(β) *the positive number* $\epsilon = \epsilon_f$ *cannot be chosen independent of* f *in* (α) *and, what is more,*

(β^*) *even if* $f(w)$ *is restricted by* $c_n \geqq 0$ *for every* n *in* (33) (*that is, even if* $f(w) \equiv f^*(w)$), *there belongs to every* $\epsilon > 0$ *some* $f(w) = f^\epsilon(w)$ *corresponding to which the function* $w(z) = w_f(z)$ *becomes singular within the circle* $|z| < 1 + \epsilon$.

(α) is a particular case of what was proved in [14] and (β) is the result of [8]. In view of the contrast between (I) and (II), the improvement of (β) to (β^*) is relevant from the point of *any* majorant of (6); in fact, (6) is its own *best majorant* if $f = f^*$. The proof of (β^*) proceeds as follows:

Let $0 < \epsilon < 1$ (eventually, $\epsilon \to 0$) and choose c_n to be $1 - \epsilon$ or $\epsilon/(2n)^2$ according as $n = 0$ or $n > 0$. Then (33) has positive coefficients, $f(w)$ is

regular for $|w|<1$ but not at $w=1$, and (19) is satisfied. It will be shown that if R is defined by

$$R=\int_0^{1-0}(f(r))^{-1}dr \tag{34}$$

(which implies that $R<\infty$), then the solution $w(z)$ of (6) and (2) must become singular at the point $z=R$. This will prove (β^*), since it will imply that $w(z)$ must become singular within the circle $|z|<1/(1-\epsilon)$ (where the denominator can be chosen arbitrarily close to 1). In fact, since $f(r)$ is positive and increasing on the interval $0\leqq r<1$, it is clear from (34) that $R<1/f(0)$, where $f(0)=c_0=1-\epsilon$.

It follows from (2) and (6), where $f(0)\neq 0$, that (for small $|z|$) the function $w=w(z)$ is the (local) inverse of the function $z=z(w)$ defined (for small $|w|$) by

$$z=\int_0^{w}(f(t))^{-1}dt. \tag{35}$$

Consider the z-map of the interval for $0\leqq w<1$. Since the function $f(w)=f^*(w)$ is positive for $0\leqq w<1$ and has a positive and increasing derivative for $0<w<1$, it is clear from (35) and (34) that the interval $0\leqq w<1$ has the *schlicht* image $0\leqq z<R$, and that, since (19) is satisfied, dz/dw stays between two positive bounds as $w\to 1$. Hence, if $w(z)$ did not become singular at $z=R$, then $z(w)$ would remain regular at $w=1$. In view of (35), this is possible only if $1/f(w)$ remains regular at $w=1$. Since $f(w)$ was chosen to be singular at $w=1$, it follows that $f(w)$ has a pole at $w=1$. But this contradicts (19).

Appendix.

In the assumptions of (i), the absolute value of the given function $f(z,w)$ is limited by 1 from *above*. If this limitation is made from *below*, the result is as follows:

(iii) *If $f(z,w)$ is regular, and satisfies the inequality $|f(z,w)|\geqq 1$, on the dicylinder $(|z|<1, |w|<1)$, then the solution $w(z)$ of (1) and (6) is regular, and satisfies the inequality $|w(z)|<1$, on the circle $|z|<|f(0,0)|^{-2}$.*

In the proof of (iii), the following lemma (*) will be needed:

(*) *If* $w = w(z)$, *where* $|z| < 1$, *is any regular function satisfying the conditions* $w(0) = 0$, $w'(0) \neq 0$, *where* $w' = dw/dz$, *and the inequality* $|w'(z)| < 1$ *for* $|z| < 1$, *then the inverse function* $z = z(w)$ *is regular, and in absolute value less than* 1, *on the circle*

$$|w| < |w'(0)|^2. \tag{36}$$

As will be seen in a moment, (*) supplies not only (iii) but also the following fact:

(iii bis) *Under the assumptions and in the notations of* (iii), *the function* $w(z)$ *is schlicht in the circle* $|z| < |f(0,0)|^{-2}$.

Similarly, (i) can now be completed as follows:

(i bis) *Under the assumptions and in the notations of* (i), *the inverse,* $z = z(w)$, *of the function* $w(z)$ *is regular and schlicht in the circle* $|w| < |f(0,0)|^2$, *provided that* $f(0,0) \neq 0$.

In fact, (i bis) is clear from (i) and (*). It is also clear that both (iii) and (iii bis) follow from (*) and (i) if z and w are interchanged and, correspondingly, (6) and (2) are written as $dz/dw = F(w,z)$ and $z(0) = 0$, where $F(w,z) = 1/f(z,w)$ and $z = z(w)$. Thus it is sufficient to verify (*).

Remark. If $f(z,w)$ is of the particular form $f(w)$, as in (ii), then (i bis), where $|f(0,0)| < 1$ by assumption, can be completed as follows:

(ii bis) *Under the assumptions and in the notations of* (ii), *the function* $w(z)$ *is schlicht in the circle* $|z| < 1$, *provided that* $f(0) \neq 0$ (*if* $f(0) = 0$, *then* $w(z) \equiv 0$).

(ii bis) is a corollary of the sharper results on the local inverse of the integral (35) which are contained in a paper written in cooperation with Dr. Hartman (*Rend. Palermo,* ser. 2, vol. 3 (1954), pp. 286-292). But (ii bis) itself is trivial. For, on the one hand, the solution $w(z)$ of (6) and (2) is regular, and satisfies $|w(z)| < 1$, for $|z| < 1$ and, on the other hand, (35) defines z as a single-valued and regular function if w and the integration path joining w to 0 are confined to any simply-connected domain which is contained in the circle $|w| < 1$ and from which the zeros of $f(w)$

(if there are any in $|w|<1$) have been excluded (by joining these zeros to a point of the circumference $|w|=1$ by cuts).

Proof of (*). If the constants which in Satz X of Landau [3], p. 473, are denoted by M, R and a are chosen to be 1, 1 and $|w'(0)|$ respectively, it follows that the assertion of (*) is certainly true if the radius, $|w'(0)|^2$, of the circle (36), claimed in (*), is replaced by

$$1+(|w'(0)|^{-2}-1)\log(1-|w'(0)|^2). \tag{37}$$

Hence (*) will be proved if it is ascertained that the value of (37), where $0<|w'(0)|<1$, exceeds $|w'(0)|^2$.

Clearly, $\log(1-|w'(0)|^2)=-(|w'(0)|^2+\frac{1}{2}|w'(0)|^4+\cdots)$, where all the higher terms, those indicated by dots, have positive coefficients. Hence the value of (37) exceeds $1+(1-|w'(0)|^{-2})|w'(0)|^2$, which is $|w'(0)|^2$.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

PART I.

[1] A. Haar, "Über Eindeutigkeit und Analytizität der Lösungen partieller Differentialgleichungen," *Atti del Congresso Internazionale dei Matematici*, vol. 3 (1928), pp. 5-10.

[2] O. Hölder, "Abschätzungen in der Theorie der Differentialgleichungen," *Schwarz Festschrift* (1914), pp. 116-132.

[3] E. Landau, "Der Picard-Schottkysche Satz und die Blochsche Konstante," *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, vol. 1926, pp. 467-474.

[4] E. Lindelöf, "Sur l'application des méthodes d'approximations successives à l'étude des intégrales réelles des équations différentielles ordinaires," *Journal de Mathématiques*, ser. 4, vol. 10 (1894), pp. 117-128.

[5] R. Lipschitz, *Differential- und Integralrechnung* (1880), pp. 506-511.

[6] M. Müller, "Über den Konvergenzbereich des Verfahrens der schrittweisen Näherungen bei gewöhnlichen Differentialgleichungen," *Mathematische Zeitschrift*, vol. 41 (1936), pp. 163-175.

[7] P. Painlevé, "Sur le calcul des intégrales d'un système différentiel par la méthode de Cauchy-Lipschitz," *Bulletin de la Societé Mathématique de France*, vol. 27 (1899), pp. 149-152.

[8] A. Wintner, "On the exact value of the bound for the regularity of solutions of ordinary differential equations," *American Journal of Mathematics*, vol. 57 (1935), pp. 539-540.

PART II.

[9] E. Landau, "Über die Blochsche Konstante und zwei verwandte Weltkonstanten," *Mathematische Zeitschrift*, vol. 30 (1929), pp. 608-634.

[10] ———, *Darstellung und Begründung einiger neuerer Ergebnisse der Funktionentheorie*, 2nd ed. (1929).

[11] E. Lindelöf, "Démonstration élémentaire de l'existence des fonctions implicites," *Bulletin des Sciences Mathématiques*, ser. 2, vol. 23 (1899), part I, pp. 68-75.

[12] H. Nakano, "Über den Konvergenzradius der Lösung einer Differentialgleichung $dy/dx = f(x, y)$," *Proceedings of the Imperial Academy of Japan*, vol. 8 (1932), pp. 29-31.

[13] I. Schur, "Über Potenzreihen, die im Innern des Einheitskreises beschränkt sind.," *Journal für die reine und angewandte Mathematik*, vol. 147 (1917), pp. 205-232 and vol. 148 (1918), pp. 122-145.

[14] A. Wintner, "On the bound of regularity of the solution of analytic differential equations of first order," *Quarterly Journal of Mathematics* (Oxford), ser. 2, vol. 5 (1954), pp. 145-149.

[15] ———, "Sur le calcul des limites de Cauchy dans la théorie des équations différentielles ordinaires," *Comptes Rendus*, vol. 242 (1956), pp. 1106-1107.

[16] ———, "On an absolute constant pertaining to Cauchy's 'principal moduli' in bounded power series," to appear in the *Mathematica Scandinavica* (1956).

ALGEBRAIC GROUPS OVER FINITE FIELDS.*

By SERGE LANG.

1. Introduction. Let k be a finite field with q elements. Let G be an algebraic group defined over k. (For the foundations of the theory of algebraic groups and homogeneous spaces, see Weil [8], [9].) If x is a point of G, we denote by $x^{(q)}$ the point obtained by raising all coordinates of x to the q-th power, i.e. by applying to x the Frobenius automorphism of the universal domain leaving k fixed. The mapping $f(x) = x^{-1}x^{(q)}$ is a rational map of G into itself. It will be shown that it is in fact surjective, and that it gives a Galois, in general non abelian unramified covering of G over itself, the Galois group being that of the left rational translations. This sort of covering is of course impossible in characteristic 0.

More generally, it will be shown that G acts on itself as a homogeneous space, under the operation $F(x, y) = x \cdot y = x^{(q)}yx^{-1}$. Using this fact we shall show that every homogeneous space H of G defined over k has a rational point. If it is principal homogeneous, then it must be biregularly equivalent to G over k, and in case G is commutative, this means that the Weil group is trivial. We also use this result to give a new proof of a result due to Chatelet [4], that if a variety V/k becomes biregularly equivalent to projective space over the algebraic closure of k, then it is already so over k itself.

Finally we consider the class field theory for our covering defined by the map $z \to z^{-1}z^{(q)}$, get a partial non abelian reciprocity law, and prove that the Artin L-series are trivial. This is used to prove the following result: Let $\mathfrak{g}$ be a subgroup of the rational points of G over k. Let H be the homogeneous space of cosets of $G \bmod \mathfrak{g}$. Then G and H have the same number of rational points.

In case the group G is commutative, then one can get a complete reciprocity law, and one can use it to derive the class field theory over a variety V having a rational map into G by means of which the abelian coverings of G by commutative groups can be pulled back in a one-one manner.

* Received March 9, 1956.

This abelian class field theory is carried out in detail elsewhere, and in this paper, we have concentrated exclusively on the non abelian aspects of the covering $z^{-1}z^{(q)}$.

2. The map $f(z) = z^{-1}z^{(q)}$. The map $f(z)$ being as above we contend that if z is a generic point of G/k then $f(z)$ is also a generic point of G/k, and in fact the extension $k(z)$ over $k(f(z))$ is separable algebraic. Namely, putting $x = f(z)$ we have $k(z) = k(x, z) = k(x, z^{(q)})$. Our contention now follows from the following elementary and well known result of field theory:

PROPOSITION 1. *Let F be a field and E/F a finitely generated extension. We assume the characteristic p is $\neq 0$. If E/F is separable algebraic, then $E^{p^\mu}F = E$ for all powers p^μ, and conversely, if $E^{p^\mu}F = E$ for some power p^μ, then E/F is separably algebraic.*

(By E^{p^μ} we denote the field obtained by raising all elements of E to the p^μ-th power.)

We have trivially for z, w in G:

$$(zw)^{(q)} = z^{(q)}w^{(q)}, \quad (z^{-1})^{(q)} = (z^{(q)})^{-1}. \tag{1}$$

Furthermore our rational map satisfies the following formalism:

$$f(zw) = f(z)^w f(w) \tag{2}$$

where y^w is defined to be $w^{-1}yw$.

The subgroup of G consisting of the elements rational over the field k_d (unique extension of k of degree d) will be denoted by G_d.

(3) An element a of G is in G_1 if and only if $f(a) = 0$. More generally, $f(z) = f(w)$ if and only if $z = aw$ for some a in G_1.

The first part of the statement is obvious. If $z = aw$, it is also obvious that $f(z) = f(w)$. Suppose $f(z) = f(w)$. Then $z^{-1}z^{(q)} = w^{-1}w^{(q)}$, whence $(zw^{-1}) = (zw^{-1})^{(q)}$. This means that zw^{-1} is rational over k, as desired.

We now see that if z is a generic point of G/k, then the extension $k(z)/k(f(z))$ is Galois, the distinct conjugates of z over $k(f(z))$ being az, with a rational over k.

The mapping $g(z) = z^{(q)}z^{-1}$ has analogous properties with respect to right translations, and we shall use them freely whenever needed.

PROPOSITION 2. *Let y be an arbitrary point of G and x a generic point over $k(y)$. Then $x^{(q)}yx^{-1}$ is a generic point of G over $k(y)$.*

Proof. Let $w = x^{(q)}yx^{-1}$. Put $K = k(y)$. Then $K(w, x) = K(w, x^{(q)})$. This implies that $K(x)$ is separable algebraic over $K(w)$, and that w has the same dimension as x over K. Hence it is a generic point of G/K.

Given two arbitrary points x and y of G, we denote by $x \cdot y$ the point $x^{(q)}yx^{-1}$. We obviously have for x, y, z arbitrary,

$$(xy) \cdot z = x \cdot (y \cdot z) \quad \text{and} \quad e \cdot x = x \tag{4}$$

To prove that G is a homogeneous space over itself with the above defined external law of composition, we need only prove the following statement:

THEOREM 1. *Given two points y and w in G, there exists a point x such that $x \cdot y = w$.*

Proof. In fact, using the associativity, it suffices to prove: Given y in G, there exists x such that $x \cdot y = e$. Let t be a generic point of G over $K = k(y)$. Then $t \cdot y$ is a generic point of G/K by Proposition 2. There is an isomorphism σ which is identity on K and maps $t^{(q)}t^{-1}$ on $t \cdot y = t^{(q)}yt^{-1}$. We can extend σ to the field $K(t)$. Let $u = t^{\sigma}$. Then

$$g(u) = g(t^{\sigma}) = g(t)^{\sigma} = t^{(q)}yt^{-1}.$$

If we put $x = u^{-1}t$, we have what we want.

COROLLARY. *The map $z \to z^{-1}z^{(q)}$ is surjective, i.e. given any y in G, there exists z such that $y = z^{-1}z^{(q)}$.*

Proof. According to the theorem, there exists z in G such that $z^{(q)}y^{-1}z^{-1} = e$. This element z does what is required.

From this corollary we see that we have indeed an unramified covering of G over itself. Given any point Q in G, there exists n points P such that $Q = P^{-1}P^{(q)}$, where n is the order of G_1. Given any one of them, all the others are simply the left rational translations of this one.

As another application of Theorem 1, we prove

THEOREM 2. *Let H/k be a homogeneous space over G. Then H has a rational point.*

Proof. We must show that there is some point u in H such that $u^{(q)} = u$. Let v be any point of H. Since H is defined over k, then $v^{(q)}$ is also in H. Since H is a homogeneous space, there exists an element x in G such that $xv^{(q)} = v$. By the corollary to Theorem 1, we can write $x = y^{-1}y^{(q)}$. From

this we see that $(yv)^{(q)} = (yv)$ and hence $u = yv$ is the element we are looking for.

We would like to point out here that Theorem 3 of [6] is a special case of the preceding result. Indeed, if a variety V/k becomes biregularly equivalent to an abelian variety over the algebraic closure of k, then V can be viewed as a principal homogeneous space over its Albanese variety A, which is known (by Chow's work) to be defined over k. (See Weil [9], Prop. 4.) However, because of the completeness of the group, we could give a direct proof, without using the Albanese variety. The proof can in fact be further simplified as follows:

We wish to prove that if a variety V over a finite field k becomes biregularly equivalent to a complete group variety over the algebraic closure of k, and hence over a finite extension k' of k, then V has a rational point over k. Over k' we can put a law of composition on V which makes it into a complete group variety (we don't even need to know it is abelian). There is a unit element e, rational over k', but not necessarily over k. Let z be a generic point of V over k. With respect to the composition law over k', we consider the point $f(z) = z^{-1}z^{(q)}$. It is a generic point of V over k' (by Proposition 1). Since V is complete, we can extend a specialization of $f(z)$ on e to a specialization of z to a point ξ on V, and then $\xi^{-1}\xi^{(q)} = e$. This point ξ satisfies $\xi = \xi^{(q)}$, and is therefore a rational point.

The statement made in our above mentioned paper that there is a rational place is a consequence of the following well known fact: *Let k be any field, and V/k any variety (say affine). Let Q be a simple point of V, rational over k. Let v be a generic point of V/k. Then there exists a place ϕ of $k(v)$ over k such that $\phi(v) = Q$ and such that ϕ is rational over k (i.e. k-valued).* One often says that Q is at the center of ϕ. Here of course, we have the additional property that the place can be chosen in such a way that the residue class field is canonically isomorphic to k itself: No irrationalities are needed in extending the specialization $v \to Q$ to a place of the function field. There exist many proofs of the above fact, and one of them runs along the following lines. The completion of the local ring of Q in $k(v)$ is isomorphic to a power series ring in r variables ($r = \dim V$) with coefficients in k, because Q is simple. There is therefore a canonical isomorphism of $k(v)$ in the quotient field of this power series ring. This quotient field itself can be embedded in the repeated power series field, which has obviously a k-valued place mapping all the variables on 0, one after the other. The restriction of this place to $k(v)$ is the one we are looking for.

We now return to the arbitrary algebraic group G over the finite field k.

Let $\sigma, \tau, \cdots$ range over the group of automorphisms of the extension k_d of k. This group (which is cyclic) operates on G_d in an obvious fashion. Referring to this operation, we show that the 1-cocycles split:

PROPOSITION 3. *Let $\{x_\sigma\}$ be a set of elements of G_d such that $x_\tau{}^\sigma x_\sigma = x_{\sigma\tau}$. Then there exists y in G_d such that $x_\sigma = y^\sigma y^{-1}$.*

Proof. We can change our indices from σ to integers mod d. According to the corollary of Theorem 2, we can write $x_1 = y^{(q)}y^{-1}$ for some y in G (we do not know yet whether it is in G_d). Then by the cocycle property, we get $x_i = y^{(q^i)}y^{-1}$. Finally, taking $i = d$ we must have

$$e = x_0 = x_d = y^{(q^d)}y^{-1}.$$

This shows that y is rational over k, because $y^{(q^d)} = y$.

As an application, we prove Chatelet's theorem:

THEOREM 3. *Let V be an abstract variety defined over k, which becomes biregularly equivalent to projective space S over the algebraic closure of k. Then V is biregularly equivalent to S over k.*

Proof. Let $T: V \to S$ be the correspondence, defined over some k_d. With Chatelet ([3], [4]) we take $x_\sigma = T^\sigma T^{-1}$, which is a birational biregular correspondence between S and itself. It is therefore projective, and is an element of the projective group G, rational over k_d. It satisfies the condition of Proposition 3, and if we let y be the projective transformation as in Proposition 3, we consider $T_1 = y^{-1}T$. Then T_1 is obviously fixed under every automorphism σ of k_d over k, and hence T_1 is defined over k.

For a proof that the only biregular correspondences of S with itself are projective, see Chow [5]. In that paper, other varieties are proved to have that property, and our theorem applies to them as well.

3. Class field theory. We shall now investigate the covering $f(x) = x^{-1}x^{(q)}$ from the point of view of class field theory. By a cycle, we shall always mean a cycle of dimension 0. Let $\mathfrak{p}$ be a prime rational cycle of G over k, and let Q be any point in it. Let P be any point such that $f(P) = Q$. All other points are of type aP, where a is rational over k. If $\mathfrak{p}$ is or degree d, then $f(P^{(q^d)}) = f(P)^{(q^d)} = Q$. Hence $P^{(q^d)}$ also lies in the inverse image of Q under f, and hence there exists a rational point a in G_1 such that $aP = P^{(q^d)}$.

We define $\pi_d(Q)$ to be the product $QQ^{(q)} \cdot \cdot \cdot Q^{(q^{d-1})}$. Then we have

$$aP = P^{(q^d)} = P\pi_d(Q). \tag{6}$$

It is clear that the point a in G_1 is completely well defined by the prime $\mathfrak{p}$, up to conjugacy in G_1. Furthermore we have

PROPOSITION 4. *Given two points Q_1 and Q_2 in G_1, let a_1 and a_2 in G_1 be any points determined by the condition*

$$a_1P_1 = P_1^{(q)} = P_1Q_1, \qquad a_2P_2 = P_2^{(q)} = P_2Q_2.$$

Then a_1 is conjugate to a_2 in G_1 if and only if Q_1 is conjugate to Q_2 in G_1.

The proof is trivial and formal, and we leave it to the reader.

We have a mapping from the primes to the conjugacy classes of G_1 defined by (6), and if the prime is of degree 1, then we obtain a 1-1 mapping of the conjugacy classes of G_1 (viewed as a set of primes of degree one) onto the conjugacy classes of G_1 (viewed as Galois group of left translations). The period of the point a is clearly equal to the period of $\pi_d(Q)$, and thus we can tell the period of the Frobenius class associated with a prime. Furthermore, we see that $\mathfrak{p}$ splits completely if and only if $\pi_d(Q) = e$.

I have not been able to determine whether the conjugacy class of a is always equal to that of Q (when Q is in G_1) and more generally to determine if a and $\pi_d(Q)$ are conjugate in G_d. In order to obtain the complete decomposition laws, we would need to know that any rational point b in G_1, conjugate to $\pi_d(Q)$ in G_d is conjugate to a in G_1. This would imply that the Frobenius class associated with $\mathfrak{p}$ in G_1 can be determined rationally. (This is the case for the full linear group.)

We now consider the L-series.

Let z be a generic points of G/k. Let

$$f_n(z) = z^{-1}z^{(q^n)} \text{ and } \pi_n(z) = zz^{(q)} \cdot \cdot \cdot z^{(q^{n-1})}$$

Then $\pi_n(f_1(z)) = f_n(z)$.

Let $E = k(z)$, $F = k(f_1(z))$, and $K = k(f_n(z))$. We have inclusions $E \supset F \supset K$, and E/F is Galois (it is the extension discussed previously). Let E_n, F_n, and K_n be the constant field extensions by k_n. Then E_n/K_n becomes Galois, with group G_n. The group G is a model of each one of our function fields, but of course in a different way each time.

Referring to these models, we shall prove that the L-series are trivial for a character not containing the identity.

Let χ be a character of G_1. Let $\mathfrak{p}$ be a prime cycle of G/k. Let $T_\mathfrak{p}$ be any one of the translations of G_1 in the Frobenius class associated with $\mathfrak{p}$ in G_1. Then the L-series are defined by

$$t\frac{d}{dt}\operatorname{Log} L(t,\chi,E/F) = \sum \chi(T_\mathfrak{p}{}^{\mu})\deg(\mathfrak{p})\, t^{\mu\deg(\mathfrak{p})}$$

the sum being taken over all primes and all $\mu \geqq 1$.

If we look at the coefficient of t^n we see that we must prove that

$$\sum_{\deg(\mathfrak{p})\mid n} \chi(T_\mathfrak{p}{}^{n/\deg(\mathfrak{p})})\deg(\mathfrak{p}) = 0.$$

This sum can be rewritten in terms of points in G_n as follows: To each point Q in G_n, we can associate a translation $T_Q{}^{(n)}$ (well defined up to conjugacy) in G_1, such that for any P in $f_1{}^{-1}(Q)$ we have $T_Q{}^{(n)}(P) = P^{(q^n)}$. It is then clear that the above sum is equal to

$$\sum_{Q\in G_n} \chi(T_Q{}^{(n)}). \tag{7}$$

For $n = 1$ we know by Proposition 4 that each class will have a representative $T_Q{}^{(1)}$ for some Q in G_1, and that this representative will occur in our sum as many times as there are element in that class. Consequently the value of our sum is the same as the value of the character taken over the sum of the conjugacy classes in the group ring of G_1. If χ does not contain the identity character, then this sum must be 0.

For arbitrary n, we note that the coefficient of t^n in our L-series $L(t,\chi,E/F)$ is by definition and formula (7) the same as the coefficient of t in $L(t,\chi,E_n/F_n)$. We have thus reduced the computation of the n-th term to the computation of the first term of an L-series over a bigger constant field. By one of the main theorems on L-series, we know that

$$L(t.\chi,E_n/F_n) = L(t,\chi^*,E_n/K_n)$$

where χ^* is the induced character. If χ does not contain the identity, then neither does χ^*. The extension E_n/K_n is now of a type analogous to that considered above for $n = 1$ (i.e. belonging to a rational map f_n). Hence the first coefficient of the L-series is equal to zero, as desired.

Let $\lambda\colon G \to H$ be a homomorphism of an algebraic group G onto an algebraic group H, defined over k, and with finite kernel. Let z be a generic point of G/k. As usual, $f_G(z) = z^{-1}z^{(q)}$. We also have an f-mapping on H, denoted by f_H. Then obviously $f_H\lambda = \lambda f_G$. Taking the degree of both sides, we see that the degree of f_H must equal that of f_G, i.e. that *G and H have*

the same number of rational points (hence the same zeta function). If kernel of λ is contained in G_1, and λ is separable, then G is Galois over H, and this again suggests that the L-series for the covering are trivial. This is indeed the case, and can be proved as follows: Put $y = \lambda(z)$. Then $E = k(z) \supset k(y) = M \supset k(z^{-1}z^{(q)}) = F$. If χ is a character for the Galois group of E/M, and does not contain the identity, then the induced character χ^* to the Galois group of E/F does not contain the identity either, and by what has been proved before, the L-series belonging to it must be trivial.

More generally, if the identity for E/M occurs with some multiplicity in χ, then the identity for E/F occurs with the same multiplicity in χ^*. (See for instance Brauer-Tate [2], where we put $\Theta =$ identity in formula (5).) From this remark we can deduce the following result.

Let H be the homogeneous space obtained from G by the cosets of a subgroup of G_1. Then we have a rational map $\lambda: G \to H$ (not necessarily a homomorphism), and the same type of field inclusion as before: $k(z) \supset k(y) \supset k(x)$. The zeta function of H, denoted by $Z_H(t)$, can be written

$$Z_H(t) = L(t, 1, E/M) = L(t, 1^*, E/F)$$

where 1 stands for the identity character on the Galois group of E/M. The identity for E/F occurs exactly once in 1^*, and hence the above L-series is equal to $Z_G(t) \cdot L(t, \chi, E/F)$, where χ is some character for E/F, which does not contain the identity, and $Z_G(t)$ is the zeta function of G. This shows that the zeta function of G and H coincide, and hence that G and H have the same number of rational points.

Knowing that the L-series are trivial, we can of course apply the formal argument given by Artin (Satz 4 of [1]) to get the density of primes in a given arithmetic progression. Let χ_i be the simple characters of G_1. Then we know that

$$\sum_{Q \in G_n} \chi_i(T_Q^{(n)}) = \begin{cases} 0 & i \neq 1 \\ q^{nr} + O(q^{n(r-1/2)}) & i = 1 \end{cases} \tag{8}$$

because for $i = 1$, we deal with the zeta function and can use the results of [7]. Let C_j be a fixed class in G_1 and T any element of C_j. Let h be the order of G_1 and h_j the number of elements in C_j. Multiplying (8) by $\chi_i(T^{-1})$ and summing, we use the orthogonality relations (see formula (3) of [1]) to get

$$\frac{h}{h_j} N(n, C_j) = q^{nr} + O(q^{n(r-1/2)}),$$

where $N(n, C_j)$ is the number of points in G_n having their $T_Q^{(n)}$ in the given class C_j. One sees trivially that to get an estimate for the number of primes, one has to divide by n.

COLUMBIA UNIVERSITY.

REFERENCES.

[1] E. Artin, "Über eine neue Art von L-Reihen," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 3 (1923), pp. 89-108.

[2] R. Brauer and J. Tate, "On the characters of finite groups," *Annals of Mathematics*, vol. 62 (1955), pp. 1-7.

[3] F. Chatelet, "Variations sur un thème de Poincaré," *Annales de l'Ecole normale superieure*, vol. 61 (1944), pp. 249-300.

[4] ———, "Les courbes de genre 1 dans un champ de Galois," *Comptes rendus des seances de l'Academie des Sciences*, vol. 224 (1947), pp. 1616-1618.

[5] W. L. Chow, "On the geometry of algebraic homogeneous spaces," *Annals of Mathematics*, vol. 50 (1949), pp. 32-67.

[6] S. Lang, "Abelian varieties over finite fields," *Proceedings of the National Academy of Sciences*, vol. 41 (1955), pp. 174-176.

[7] ——— and A. Weil, "Number of points of varieties in finite fields," *American Journal of Mathematics*, vol. 76 (1954), pp. 819-827.

[8] A. Weil, "Algebraic groups of transformations," *American Journal of Mathematics*, vol. 77 (1955), pp. 355-391.

[9] ———, "On algebraic groups and homogeneous spaces," *American Journal of Mathematics*, vol. 77 (July 1955), pp. 493-512.

REPRESENTATIONS OF SEMISIMPLE LIE GROUPS VI.*

Integrable and Square-Integrable Representations.

By HARISH-CHANDRA.

---

1. Introduction. Let G be a connected semisimple Lie group and let Z denote its center. Then if π is an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$ we can find a unitary character η_π of Z such that $\pi(z) = \eta_\pi(z)\pi(1)$ for $z \in Z$. Let $x \to x^*$ denote the natural mapping of G on $G^* = G/Z$. If $\phi, \psi \in \mathfrak{H}$, it is clear that $(\phi, \pi(x)\psi)\,(x \in G)$ depends only on x^*. Let dx^* denote the Haar measure on G^*. We shall say that π is square-integrable if there exists an element $\psi_0 \neq 0$ in such that $\int_{G^*} |(\psi_0, \pi(x)\psi_0)|^2\, dx^* < \infty$. Similarly π is said to be integrable if $\int_{G^*} |(\psi_0, \pi(x)\psi_0)|\, dx^* < \infty$ for some $\psi_0 \neq 0$ in $\mathfrak{H}$. In this paper we intend to study in detail some examples of such representations.

Let π be a square-integrable representation of G. Then, as shown by Godement [4(b)], the Schur orthogonality relations hold for π and therefore there exists a positive constant d_π such that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2\, dx^* = d_\pi^{-1} |\phi|^2 |\psi|^2$$

for all $\phi, \psi \in \mathfrak{H}$. Naturally d_π depends on the normalization of the measure dx^* but once this has been fixed, d_π can be considered as a function of π. In analogy with the case of compact groups we call d_π the formal degree of π. If Z is finite and ω is the equivalence class of π, d_π is also equal to the mass of ω with respect to the Plancherel measure (see Section 5) just as in the compact case. Moreover for compact semisimple groups Weyl [11(a)] has given a formula for the degree of an irreducible representation in terms of its "highest weight." We shall se that substantially the same formula holds for the formal degree d_π under suitable conditions. This is the principal result of this paper. It can be verified immediately in the case of the 2×2 real unimodular group by looking at the results obtained

---

* Received September 27, 1955.

by Bargmann [1, p. 634] by direct computation. This very simple case is, in a sense, fundamental and much of our argument will depend on the properties of this three-dimensional group.

This paper is divided into two parts. In Part I we obtain the Schur orthogonality relations and establish the connection between the formal degree and the Plancherel measure. We also obtain a formula for the character of a square-integrable representation which is quite similar to the corresponding formula in the compact case. Although the Schur orthogonality relations are now new, Godement's proof of them [4(b)] does not cover the case when Z is infinite. (However it could perhaps be modified to include this case as well.) At any rate our method is quite different and it seems worthwhile to present it briefly even at the risk of some overlap with earlier work especially since the infinite case mentioned above is particularly important for us.

Part II is devoted to the proof of the analogue of Weyl's formula. This proof depends on a detailed comparison at each step between the compact and the non-compact cases and the entire argument is based on Lemma 22. In order to make this comparison we need some considerable algebraic preparation which consists of an intensive study of the root-structure of certain types of semisimple Lie algebras. As an incidental outcome of this study, we get in Section 7 a new proof of a theorem of E. Cartan [2(c), p. 145)] on the boundedness of certain complex domains. This proof, unlike that of Cartan, does not depend on the classification of simple groups.

The last two sections contain the proof of Lemma 22. Here I follow closely a method due to Weyl [11(a)] and Cartan [2(a)] and although the proof is somewhat long no new ideas are involved.

The results of this paper had been announced in a short note [5(d)].

Part I.

2. Preliminary lemmas. Let G be a connected semisimple Lie group and let $\mathfrak{g}_0$ denote its Lie algebra over the field R of real numbers. Define $\mathfrak{k}_0$ as in [5(b)] and let $\mathfrak{c}_0$ be the center and $\mathfrak{k}_0' = [\mathfrak{k}_0, \mathfrak{k}_0]$ the derived algebra of $\mathfrak{k}_0$. Let K, K' and D denote the analytic subgroups of G corresponding to $\mathfrak{k}_0$, $\mathfrak{k}_0'$ and $\mathfrak{c}_0$ respectively. We consider the space $C_c(G)$ of all (complex-valued) continuous functions on G which vanish outside a compact set. For any two functions $f, g \in C_c(G)$ we define their convolution $f * g$ by

$$(f * g)(x) = \int_{G} f(y) g(y^{-1}x)\, dy \qquad (x \in G)$$

where dy is the Haar measure on G. Under this operation $C_c(G)$ becomes an associative algebra. Let Ω denote the set of all equivalence classes of finite-dimensional irreducible representations of K and let $\xi_{\mathfrak{D}}$ be the character (on K) of any class $\mathfrak{D} \in \Omega$. Choose a base $\Gamma_1, \cdots, \Gamma_r$ for $\mathfrak{c}_0$ over R such that $\exp \Gamma_i \in D \cap Z$, $1 \leqq i \leqq r$. (Z is the center of G.) This is possible since $D/D \cap Z$ is compact (see Mostow [9]). Let $\mathfrak{c}_1$ be the subset of $\mathfrak{c}_0$ consisting of all elements of the form $t_1\Gamma_1 + \cdots + t_r\Gamma_r$ ($t_i \in R$) with $|t_i| \leqq \frac{1}{2}$, $1 \leqq i \leqq r$. Then $K_0 = K'(\exp \mathfrak{c}_1)$ is a compact subset of K. For any $f \in C_c(G)$ and $\mathfrak{D} \in \Omega$ we define two functions ${}_{\mathfrak{D}}f$ and $f_{\mathfrak{D}}$ in $C_c(G)$ as follows:

$$ {}_{\mathfrak{D}}f(x) = d(\mathfrak{D}) \int_{K_0} \xi_{\mathfrak{D}}(u) f(ux)\, du, \qquad f_{\mathfrak{D}}(x) = d(\mathfrak{D}) \int_{K_0} f(xu) \xi_{\mathfrak{D}}(u)\, du $$

($x \in G$). Here du is the Haar measure on K normalized in such a way that $\int_{K_0} du = 1$. Also $d(\mathfrak{D})$ is the degree of any representation in $\mathfrak{D}$. Let $L(\mathfrak{D})$ ($\mathfrak{D} \in \Omega$) denote the subspace of $C_c(G)$ consisting of all functions of the form ${}_{\mathfrak{D}}f$ ($f \in C_c(G)$). Since ${}_{\mathfrak{D}}(f * g) = ({}_{\mathfrak{D}}f) * g$ ($f, g \in C_c(G)$ it follows that $L(\mathfrak{D})$ is a right ideal in $C_c(G)$.

Let π be a quasi-simple irreducible representation (see [5(b)] of G on a Banach space $\mathfrak{H}$. For any $\mathfrak{D} \in \Omega$, let $\mathfrak{H}_{\mathfrak{D}}$ denote the subspace of $\mathfrak{H}$ consisting of all those elements which transform under $\pi(K)$ according to $\mathfrak{D}$. Let Ω_π be the set of all $\mathfrak{D} \in \Omega$ such that $\mathfrak{H}_{\mathfrak{D}} \neq 0$. Then $L_\pi = \sum_{\mathfrak{D} \in \Omega_\pi} L(\mathfrak{D})$ is a subalgebra of $C_c(G)$. Since π is quasi-simple there exists a character η_π of Z such that $\pi(z) = \eta_\pi(z)\pi(1)$ ($z \in Z$). We shall call η_π *the central character of* π. For any $f \in C_c(G)$ let $\pi(f)$ denote the operator $\int_G f(x)\pi(x)\, dx$. Then $f \to \pi(f)$ defines a representation of $C_c(G)$ on $\mathfrak{H}$.

LEMMA 1. *The space* $\mathfrak{H}_0 = \sum_{\mathfrak{D} \in \Omega} \mathfrak{H}_{\mathfrak{D}}$ *is invariant and* (*algebraically*) *irreducible under* $\pi(L_\pi)$.

Let $E_{\mathfrak{D}}$ ($\mathfrak{D} \in \Omega$) denote the canonical projection [5(b), p. 225] of $\mathfrak{H}$ on $\mathfrak{H}_{\mathfrak{D}}$. Then if $\mathfrak{D} \in \Omega_\pi$, one proves easily (see [5(c), p. 249]) that

$$ E_{\mathfrak{D}} = d(\mathfrak{D}) \int_{K_0} \xi_{\mathfrak{D}}(u^{-1}) \pi(u)\, du $$

and therefore

$$ E_{\mathfrak{D}} \pi(f) = \pi({}_{\mathfrak{D}}f) \qquad (f \in C_c(G)). $$

On the other hand if $\mathfrak{D} \notin \Omega_\pi$, $E_{\mathfrak{D}} = 0$. Hence it is clear that if $f \in L_\pi$, $\pi(f)$ maps $\mathfrak{H}$ into $\mathfrak{H}_0$. Let ψ_0 be any nonzero element in $\mathfrak{H}_0$. In order to prove

the irreducibility of $\mathfrak{H}_0$ under $\pi(L_\pi)$ it would be enough to show that $\mathfrak{H}_\mathfrak{D} \subset \pi(L_\pi)\psi_0$ for all $\mathfrak{D} \in \Omega_\pi$. Suppose then that this is false for some $\mathfrak{D}$. Put $U = \mathfrak{H}_\mathfrak{D} \cap \pi(L_\pi)\psi_0$. Then $U \neq \mathfrak{H}_\mathfrak{D}$ and since $\dim \mathfrak{H}_\mathfrak{D}$ is finite [5(b)], there exists a linear function $\alpha \neq 0$ on $\mathfrak{H}_\mathfrak{D}$ which vanishes identically on U. Extend α to a continuous linear function on $\mathfrak{H}$ by setting $\alpha(\psi) = \alpha(E_\mathfrak{D}\psi)$ $(\psi \in \mathfrak{H})$. Since $\psi_0 \neq 0$ and π is an irreducible representation of G, it is obvious that the continuous function $\alpha(\pi(x)\psi_0)$ $(x \in G)$ cannot be everywhere zero on G. Hence we can choose $f \in C_c(G)$ such that

$$\int f(x)\alpha(\pi(x)\psi_0)\,dx \neq 0.$$

Then

$$\alpha(\pi({}_\mathfrak{D}f)\psi_0) = \alpha(E_\mathfrak{D}\pi(f)\psi_0) = \alpha(\pi(f)\psi_0) = \int f(x)\alpha(\pi(x)\psi_0)\,dx \neq 0.$$

But since ${}_\mathfrak{D}f \in L(D)$, $\pi({}_\mathfrak{D}f)\psi_0 \in U$ and therefore $\alpha(\pi({}_\mathfrak{D}f)\psi_0) = 0$. This contradiction proves the lemma.

COROLLARY. *For any $\mathfrak{D} \in \Omega_\pi$, $\mathfrak{H}_\mathfrak{D}$ is invariant and irreducible under $\pi(L(\mathfrak{D}))$ and the corresponding representation of $L(\mathfrak{D})$ on $\mathfrak{H}_\mathfrak{D}$ determines π completely up to infinitesimal equivalence* [5(b), p. 230].

Since $E_\mathfrak{D}\pi(f) = \pi({}_\mathfrak{D}f)$ $(f \in C_c(G))$ it is obvious that $\mathfrak{H}_\mathfrak{D}$ is invariant under $\pi(L(\mathfrak{D}))$. Let ψ_0 and ψ be two elements in $\mathfrak{H}_\mathfrak{D}$ and suppose $\psi_0 \neq 0$. Then it follows from the above lemma that $\psi = \pi(f)\psi_0$ for some $f \in L_\pi$. But since $\psi \in \mathfrak{H}_\mathfrak{D}$,

$$\psi = E_\mathfrak{D}\psi = \pi({}_\mathfrak{D}f)\psi_0.$$

However ${}_\mathfrak{D}f \in L(\mathfrak{D})$ and so the irreducibility is proved.

Let $\phi(x)$ $(x \in G)$ denote the trace of the restriction of $E_\mathfrak{D}\pi(x)E_\mathfrak{D}$ on $\mathfrak{H}_\mathfrak{D}$. Then if $f \in C_c(G)$, $\int f(x)\phi(x)\,dx$ is the trace of the restriction of $E_\mathfrak{D}\pi(f)E_\mathfrak{D}$ on $\mathfrak{H}_\mathfrak{D}$. But since $E_\mathfrak{D}\pi(f) = \pi({}_\mathfrak{D}f)$, it follows that

$$\int f(x)\phi(x)\,dx = \int {}_\mathfrak{D}f(x)\phi(x)\,dx.$$

Hence if ν denotes the representation of $L(\mathfrak{D})$ on $\mathfrak{H}_\mathfrak{D}$

$$\int f(x)\phi(x)\,dx = \mathrm{Sp}(\nu({}_\mathfrak{D}f)) \qquad (f \in C_c(G)).$$

ϕ being a continuous function, it is now obvious that the knowledge of the trace of ν determines it completely. From this our assertion follows (see [5(c), p. 235]).

LEMMA 2. *Let $\psi \neq 0$ be an element in $\mathfrak{H}$. Then $E_{\mathfrak{D}}\psi \neq 0$ for some $\mathfrak{D} \in \Omega_\pi$.*

Let $C_c^\infty(G)$ be the set of all functions $f \in C_c(G)$ which are everywhere indefinitely differentiable. Choose a neighborhood V of 1 in G such that $|\pi(x)\psi - \psi| \leqq \frac{1}{2}|\psi|$ for $x \in V$. Since K_0 is compact, we can find another such neighborhood V' with the property that $uV'u^{-1} \subset V$ for $u \in K_0$. Select a real-valued function $g \in C_c^\infty(G)$ such that $g \geqq 0$, $\int g(x)dx = 1$ and $g = 0$ outside V'. Then if

$$f(x) = \int_{K_0} g(uxu^{-1})du \qquad (x \in G)$$

it is obvious that $f \in C_c^\infty(G)$, $f \geqq 0$, $\int f(x)dx = 1$ and $f = 0$ outside V. Moreover since $K = K_0Z$, one proves easily that $f(ux) = f(xu)$ $(u \in K, x \in G)$. Therefore $E_{\mathfrak{D}}\pi(f) = \pi(f)E_{\mathfrak{D}}$ and

$$|\pi(f)\psi - \psi| = \left|\int f(x)(\pi(x)\psi - \psi)dx\right| \leqq \int f(x)|\pi(x)\psi - \psi|\,dx \leqq \tfrac{1}{2}|\psi|.$$

Hence $\pi(f)\psi \neq 0$ and so it follows from Lemma 3 of [5(c)] that $\pi(f)E_{\mathfrak{D}}\psi = E_{\mathfrak{D}}\pi(f)\psi \neq 0$ for some $\mathfrak{D} \in \Omega_\pi$. This proves that $E_{\mathfrak{D}}\psi \neq 0$.

COROLLARY. *Let π' be another quasi-simple irreducible representation of G on a Banach space $\mathfrak{H}'$ and let $\psi' \neq 0$ be an element in $\mathfrak{H}'$. Suppose $\pi'(f)\psi' = 0$ whenever $\pi(f) = 0$ $(f \in C_c(G))$. Then π and π' are infinitesimally equivalent.*

Let η be the central character of π. For any $z \in Z$ and $f \in C_c(G)$, define a function ${}_zf \in C_c(G)$ by ${}_zf(x) = f(z^{-1}x)$ $(x \in G)$. Then $\pi({}_zf) = \eta(z)\pi(f)$ and therefore

$$\pi'(z)\pi'(f)\psi' = \pi'({}_zf)\psi' = \eta(z)\pi'(f)\psi' \qquad (z \in Z, f \in C_c(G)).$$

Since elements of the form $\pi(f)\psi'$ $(f \in C_c(G))$ are dense in $\mathfrak{H}'$, we conclude that η is also the central character of π'.

Now $Z \subset K$ and therefore, by Schur's lemma, there exists, for each $\mathfrak{D} \in \Omega$, a character $\eta_{\mathfrak{D}}$ of Z such that $\sigma(z) = \eta_{\mathfrak{D}}(z)\sigma(1)$ $(z \in Z)$ for any representation σ in $\mathfrak{D}$. Let Ω_0 be the set of all those $\mathfrak{D}$ for which $\eta_{\mathfrak{D}} = \eta$. Then it is obvious that $\Omega_\pi \cup \Omega_{\pi'} \subset \Omega_0$ and if $E_{\mathfrak{D}}'$ is the canonical projection of $\mathfrak{H}'$ and $\mathfrak{H}_{\mathfrak{D}}'$,

$$E_{\mathfrak{D}} = d(\mathfrak{D})\int_{K_0} \xi_{\mathfrak{D}}(u^{-1})\pi(u)du, \qquad E_{\mathfrak{D}}' = d(\mathfrak{D})\int_{K_0} \xi_{\mathfrak{D}}(u^{-1})\pi'(u)du$$

for all $\mathfrak{D} \in \Omega_0$ (see [5(c), p. 249]). Now, by the above lemma, we can choose $\mathfrak{D}_0 \in \Omega_{\pi'}$ such that $E_{\mathfrak{D}_0}'\psi' \neq 0$. Then $\pi'(f)E_{\mathfrak{D}_0}'\psi' \neq 0$ for some $f \in C_c(G)$. This implies that $\pi'(f_{\mathfrak{D}_0})\psi' \neq 0$ and therefore $\pi(f)E_{\mathfrak{D}_0} = \pi(f_{\mathfrak{D}_0}) \neq 0$. Hence $E_{\mathfrak{D}_0} \neq 0$ and so $\mathfrak{D}_0 \in \Omega_\pi \cap \Omega_{\pi'}$. Let ν and ν' be the corresponding representations of $L(\mathfrak{D}_0)$ on $\mathfrak{H}_{\mathfrak{D}_0}$ and $\mathfrak{H}_{\mathfrak{D}_0}'$ respectively. In view of the corollary to Lemma 1, it is sufficient to prove that ν and ν' are equivalent. But since they are both irreducible and finite-dimensional it would be enough to show that they have the same kernel. Let $\mathfrak{M}$ be the vernel of ν in $L(\mathfrak{D}_0)$. Then it is a maximal two-sided ideal in $L(\mathfrak{D}_0)$. If $g \in \mathfrak{M}$, $\pi(g_{\mathfrak{D}_0}) = \pi(g)E_{\mathfrak{D}_0} = 0$ and therefore $\pi'(g)E_{\mathfrak{D}_0}'\psi' = \pi'(g_{\mathfrak{D}_0})\psi' = 0$. This shows that $\nu'(\mathfrak{M})E_{\mathfrak{D}_0}'\psi' = 0$. Since $E_{\mathfrak{D}_0}'\psi' \neq 0$ and ν' is irreducible, it follows that the kernel of ν' must contain $\mathfrak{M}$ and therefore coincide with it. This completes the proof.

Let Z_0 be a fixed subgroup of Z such that Z/Z_0 is finite and let $x \to x^*$ denote the natural mapping of G and $G^* = G/Z_0$. Suppose $\alpha_1, \cdots, \alpha_r$ is a finite set of continuous linear functions on $\mathfrak{H}$ and $\psi_1, \cdots, \psi_r$ are certain given elements in $\mathfrak{H}$. Put

$$f(x) = \alpha_1(\pi(x)\psi_1) + \cdots + \alpha_r(\pi(x)\psi_r) \qquad (x \in G).$$

Then if the central character of π is unitary, it is obvious that $|f(x)|$ depends only on on x^*. Let dx^* denote the Haar measure on G^*.

LEMMA 3. *Assume that the function*

$$f(x) = \alpha_1(\pi(x)\psi_1) + \cdots + \alpha_r(\pi(x)\psi_r) \qquad (x \in G)$$

is not identically zero on G. Then if the central character of π is unitary and $\int_{G^*} |f(x)|^2 dx^* < \infty$, π is infinitesimally equivalent to an irreducible unitary repersentation of G on a Hilbert space.

Let η_π denote the central character of π. Then $f(xz) = \eta_\pi(z)f(x)$ $(x \in G, z \in Z)$ and therefore it follows easily from the Peter-Weyl theorem for the compact group $K/D \cap Z$ that

$$\int_{K_0} |f(xu)|^2 du = \sum_{\mathfrak{D} \in \Omega_\pi} \int_{K_0} |f_{\mathfrak{D}}(xu)|^2 du.$$

Since $f \neq 0$, we can choose $\mathfrak{D}_0 \in \Omega_\pi$ such that $f_{\mathfrak{D}_0} \neq 0$. Then it is clear that

$$f_{\mathfrak{D}_0}(x) = \sum_{i=1}^{r} \alpha_i(\pi(x)E_{\mathfrak{D}_0}\psi_i)$$

and since

$$\int_{K_0} |f_{\mathfrak{D}_0}(xu)|^2 du \leqq \int_{K_0} |f(xu)|^2 du,$$

it follows that

$$\int |f_{\mathfrak{D}_0}(x)|^2\,dx^* \leqq \int |f(x)|^2\,dx^* < \infty.$$

Hence if we replace ψ_i by $E_{\mathfrak{D}_0}\psi_i$ and f by $f_{\mathfrak{D}_0}$, our problem is reduced to the case when the given elements of $\mathfrak{H}$ all lie in $\mathfrak{H}_{\mathfrak{D}_0}$. Moreover it is obvious that without loss of generality we may assume that $\psi_1, \cdots, \psi_r$ is a base for $\mathfrak{H}_{\mathfrak{D}_0}$ over the field C of complex numbers. Since $L(\mathfrak{D}_0)$ is irreducible under $\pi(L(\mathfrak{D}_0))$ (Corollary to Lemma 1), it follows from Burnside's Theorem that

$$\pi(g_j)\psi_k = \delta_{jk}\psi_k \qquad 1 \leqq j, k \leqq r$$

for suitable elements $g_j \in L(\mathfrak{D}_0)$. ($\delta_{jk}$ is the Kronecker symbol.) Since $f \neq 0$ we may assume that $f_1(x) = \alpha_1(\pi(x)\psi_1)$ is not identically zero. Then

$$f_1(x) = \sum_{i=1}^{r} \alpha_i(\pi(x)\pi(g_1)\psi_i) = \int_G f(xy)g_1(y)\,dy$$

and therefore

$$|f_1(x)|^2 \leqq \int_\omega |f(xy)|^2\,dy \int_G |g_1(y)|^2\,dy$$

where ω is some compact subset of G outside which g_1 is zero. This shows that $\int |f_1(x)|^2\,dx^* < \infty$ and since $f_1 \neq 0$, our problem is now reduced to the case when

$$f(x) = \alpha(\pi(x)\psi).$$

Here α is a continuous linear function on $\mathfrak{H}$ and $\psi \in \mathfrak{H}_{\mathfrak{D}_0}$.

Let U' be the Hilbert space consisting of all measurable functions g on G such that (1) $g(xz) = g(x)\eta_\pi(z)$ $(x \in G, z \in Z)$ and (2) $\int_{G^*} |g(x)|^2 dx^* < \infty$. We define a representation σ' of G on U' by $(\sigma'(y)g)(x) = g(xy)$ $(x, y \in G, g \in U')$. It is obvious that f lies in U'. Let U be the smallest closed subspace of U' containing f which is invariant under $\sigma'(G)$. We denote by σ the representation of G defined on U under σ'. It is clear that σ is unitary. We shall now show that π is infinitesimally equivalent to σ.

For any $\mathfrak{D} \in \Omega$ let $F_{\mathfrak{D}}$ denote the corresponding canonical projection in U. Also, we denote the operator $\int g(x)\sigma(x)\,dx$ $(g \in C_c(G))$ by $\sigma(g)$. Since $f \in U_{\mathfrak{D}_0}$, it follows that $F_{\mathfrak{D}_0} \neq 0$ and therefore $F_{\mathfrak{D}_0}\sigma(g) = \sigma(_{\mathfrak{D}_0}g)$. Moreover it is obvious that elements of the form $\sigma(g)f$ $(g \in C_c(G))$ are dense in U. Therefore $\sigma(L(\mathfrak{D}_0))f$ is dense in $U_{\mathfrak{D}_0} = F_{\mathfrak{D}_0}U$. Now suppose $\pi(g)\psi = 0$ $(g \in C_c(G))$. Then

$$0 = \alpha(\pi(x)\pi(g)\psi) = \int \alpha(\pi(xy)\psi)g(y)dy = \int f(xy)g(y)dy$$

and therefore $\sigma(g)f = 0$. Hence if we can prove that σ is irreducible (and therefore also quasi-simple (see [10(b)])), our assertion would follow from the Corollary to Lemma 2. However, in view of the above remarks, we can define a linear mapping A of $\mathfrak{H}_{\mathfrak{D}_0}$ onto $\sigma(L(\mathfrak{D}_0))f$ such that

$$A(\pi(g)\psi) = \sigma(g)f \qquad (g \in L(\mathfrak{D}_0)).$$

Then $\dim \sigma(L(\mathfrak{D}_0))f \leqq \dim \mathfrak{H}_{\mathfrak{D}_0} < \infty$ and therefore since $\sigma(L(\mathfrak{D}_0))f$ is dense in $U_{\mathfrak{D}_0}$, $U_{\mathfrak{D}_0} = \sigma(L(\mathfrak{D}_0))f$. But $\mathfrak{H}_{\mathfrak{D}_0}$ is irreducible under $\pi(L(\mathfrak{D}_0))$ and so it follows from the existence of A that the same is true for $U_{\mathfrak{D}_0}$ under $\sigma(L(\mathfrak{D}_0))$. Now suppose V is a closed subspace of U which is invariant under $\sigma(G)$ and let W be the orthogonal complement of V in U. In view of the above irreducibility either V or W must contain $U_{\mathfrak{D}_0}$. Suppose $V \supset U_{\mathfrak{D}_0}$. Then $f \in V$ and therefore $V = U$ from the definition of U. Similarly if $W \supset U_{\mathfrak{D}_0}$, $W = U$. This proves that σ is an irreducible representation and so our lemma follows.

COROLLARY.[1] *Let π be an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$. Suppose there exist two elements $\phi_0 \neq 0$, $\psi_0 \neq 0$ in $\mathfrak{H}$ such that*

$$\int_{G^*} |(\phi_0, \pi(x)\psi_0)|^2 dx^* < \infty.$$

Then there exists a positive real number d_π such that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 dx^* = d_\pi^{-1} |\phi|^2 |\psi|^2$$

for all ϕ, ψ in $\mathfrak{H}$.

Let V and W respectively be the subspaces of $\mathfrak{H}$ consisting of all elements of the form $\pi(f)\phi_0$ and $\pi(f)\psi_0$ $(f \in C_c(G))$. Then V and W are both dense in $\mathfrak{H}$. Since π is irreducible and unitary, it is quasi-simple [10(b)] and therefore $\dim \mathfrak{H}_{\mathfrak{D}} < \infty$ $(\mathfrak{D} \in \Omega)$. But if $\mathfrak{D} \in \Omega_\pi$, $E_{\mathfrak{D}}\pi(f) = \pi(_{\mathfrak{D}}f)$ $(f \in C_c(G))$ and hence $E_{\mathfrak{D}}V \subset V$. V being dense in $\mathfrak{H}$, it follows that $E_{\mathfrak{D}}V$ is dense in $\mathfrak{H}_{\mathfrak{D}}$ and therefore $E_{\mathfrak{D}}V = \mathfrak{H}_{\mathfrak{D}}$. This shows that $\mathfrak{H}_0 = \sum_{\mathfrak{D}} \mathfrak{H}_{\mathfrak{D}} \subset V$. Similarly $\mathfrak{H}_0 \subset W$.

If $f, g \in C_c(G)$,

$$(\pi(f)\psi_0, \pi(x)\pi(g)\psi_0) = \int\int f(y)(\phi_0, \pi(y^{-1}xz)\psi_0)g(z)dydz$$

[1] See Godement [4(b)].

and therefore it is obvious that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 \, dx < \infty$$

for $\phi \in V$ and $\psi \in W$. Hence without loss of generality we may assume that $\phi_0, \psi_0 \in \mathfrak{H}_{\mathfrak{D}_0}$ for some $\mathfrak{D}_0 \in \Omega_\pi$. Let $\phi \neq 0$ be any element in V. Put $f_\phi(x) = (\phi, \pi(x)\psi_0)$ $(x \in G)$ and define U' and σ' as in the proof of Lemma 3. Let U_ϕ be the smallest closed subspace of U' containing f_ϕ which is invariant under $\sigma'(G)$ and let σ_ϕ denote the corresponding representation of G on U_ϕ. Then as we have seen during the proof of Lemma 3, σ_ϕ is irreducible and quasi-simple. Also it is obvious that $\sigma_\phi(h)f_\phi$ $(h \in C_c(G))$ is the function $x \to (\phi, \pi(x)\pi(h)\psi_0)$ $(x \in G)$. Therefore it follows from the Corollary to Lemma 2 that π and σ_ϕ are infinitesimally equivalent. Since they are both unitary they are equivalent [5(b), Theorem 8] and so there exists a unitary mapping B_ϕ of $\mathfrak{H}$ onto U_ϕ such that $B_\phi \pi(x) = \sigma_\phi(x) B_\phi$ $(x \in G)$. Moreover in view of what we have said above, there exists a linear mapping A_ϕ of W into U_ϕ such that

$$A_\phi \pi(h)\psi_0 = \sigma_\phi(h) f_\phi \qquad (h \in C_c(G)).$$

Put $C_\phi = B_\phi^{-1} A_\phi$. Then C_ϕ is a linear mapping of W into $\mathfrak{H}$ and $\pi(h)C_\phi = C_\phi \pi(h)$ $(h \in C_c(G))$. This holds in particular if $h \in L(\mathfrak{D})$ $(\mathfrak{D} \in \Omega_\pi)$ and therefore C_ϕ maps $\mathfrak{H}_\mathfrak{D}$ into itself. Now if we apply Schur's lemma to the finite-dimensional irreducible representation of $L(\mathfrak{D})$ on $\mathfrak{H}_\mathfrak{D}$, we can conclude from Lemma 1 that C_ϕ must be a scalar multiple of the identity on $\mathfrak{H}_0$. So there exists a complex number c_ϕ such that

$$\| \sigma_\phi(h) f_\phi \| = \| A_\phi \pi(h)\psi_0 \| = |c_\phi| \, \| B_\phi \pi(h)\psi_0 \| = |c_\phi| \, |\pi(h)\psi_0|$$

if $h \in L_\pi$. (Here $\| \ \|$ denotes the norm in U_ϕ.) But since

$$\| \sigma_\phi(h) f_\phi \|^2 = \int_{G^*} |(\phi, \pi(x)\pi(h)\psi_0)|^2 \, dx^*,$$

it follows that

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 \, dx^* = |c_\phi|^2 \, |\psi|^2$$

for all $\psi \in \mathfrak{H}_0$. If $\phi = 0$, we put $c_\phi = 0$ so that the above relation continues to hold in that case as well. Now suppose ϕ, ψ lie in $\mathfrak{H}_0 \subset V \cap W$. Then

$$\int |(\phi, \pi(x)\psi)|^2 \, dx^* = \int |(\psi, \pi(x)\phi)|^2 \, dx^*$$

and therefore

$$|c_\phi|^2 \, |\psi|^2 = |c_\psi|^2 \, |\phi|^2.$$

Hence we can find a real number c such that $|c_\phi|^2 = c|\phi|^2$ for $\phi \in \mathfrak{H}_0$ and therefore

$$\int |(\phi, \pi(x)\psi)|^2 \, dx^* = c|\phi|^2|\psi|^2 \qquad (\phi, \psi \in \mathfrak{H}_0).$$

Since $\mathfrak{H}_0$ is dense in $\mathfrak{H}$, it is obvious that c is positive and an elementary argument shows that the above relation continues to hold for all $\phi, \psi \in \mathfrak{H}$. Now if we put $d_\pi = c^{-1}$ we get the assertion of the Corollary.

3. The Schur orthogonality relations.

DEFINITION. *Let π be an irreducible unitary representation of G on a Hilbert space $\mathfrak{H}$. We say that π is square-integrable, if there exist two elements $\phi_0 \neq 0$, $\psi_0 \neq 0$ in $\mathfrak{H}$ such that*

$$\int_{G^*} |(\phi_0, \pi(x)\psi_0)|^2 \, dx^* < \infty.$$

Similarly we say that π is integrable if

$$\int_{G^*} |(\phi_0, \pi(x)\psi_0)| \, dx^* < \infty$$

for some nonzero elements ϕ_0, ψ_0 in $\mathfrak{H}$.

It is obvious that the above definitions do not depend on the choice of the subgroup Z_0 so long as Z/Z_0 is finite. We have seen above (Corollary to Lemma 3) that if π is square-integrable

$$\int_{G^*} |(\phi, \pi(x)\psi)|^2 \, dx < \infty$$

for all $\phi, \psi \in \mathfrak{H}$. In analogy with the case of compact groups, we shall call the number d_π (of the Corollary to Lemma 3) the *formal degree* of π. Naturally d_π depends on the choice of Z_0 and the normalization of the Haar measure of G^*. However once these have been fixed, it is obvious that two equivalent square-integrable representations have the same formal degree.

The situation for integrable representations is somewhat similar.

LEMMA 4. *Let π be an integrable representation of G on $\mathfrak{H}$. Then if $\mathfrak{H}_0 = \sum_{\mathfrak{D} \in \Omega} \mathfrak{H}_\mathfrak{D}$,*

$$\int |(\phi, \pi(z)\psi)| \, dx^* < \infty$$

for all ϕ, ψ in $\mathfrak{H}_0$.

Choose nonzero elements ϕ_0, ψ_0 in $\mathfrak{H}$ such that

$$\int |(\phi_0, \pi(x)\psi_0)|\, dx^* < \infty.$$

Then if $g, h \in C_c(G)$, it is easy to verify that the function $|(\pi(g)\phi_0, \pi(x)\pi(h)\psi_0)|$ is integrable on G^*. Let V and W be the set of all elements of the form $\pi(g)\phi_0$ and $\pi(g)\psi_0$ $(g \in C_c(G))$ respectively. Then, as we have seen during the proof of the Corollary to Lemma 3, $\mathfrak{H}_0 \subset V \cap W$ and therefore our assertion follows.

Let π and π' be two square-integrable representations of G on the Hilbert spaces $\mathfrak{H}$ and $\mathfrak{H}'$ respectively. Then if their central characters coincide on Z_0, it is obvious that[2] $(\phi, \pi(x)\psi)\ \mathrm{conj}(\phi', \pi'(x)\psi')$ may be regarded as a function of x^* on G^* $(\phi, \psi \in \mathfrak{H}; \phi', \psi' \in \mathfrak{H}'; x \in G)$.

THEOREM 1 (The Schur orthogonality relations[1]). *If π and π' are not equivalent*

$$\int_{G^*} (\phi, \pi(x)\psi)\,\mathrm{conj}(\phi', \pi'(x)\psi')\, dx^* = 0$$

for all $\phi, \psi \in \mathfrak{H}$ and $\phi', \psi' \in \mathfrak{H}'$. On the other hand if the two representations are equivalent under a unitary mapping U of $\mathfrak{H}$ onto $\mathfrak{H}'$,

$$\int_{G^*} (\phi, \pi(x)\psi)\,\mathrm{conj}(\phi', \pi'(x)\psi')\, dx^* = d_\pi^{-1}(U\phi, \phi')\,(\psi', U\psi)$$

$(\phi, \psi \in \mathfrak{H}; \phi' \psi' \in \mathfrak{H}')$ where d_π is the formal degree of π.

Let $d_{\pi'}$ denote the formal degree of π'. Then it is obvious from the Corollary to Lemma 3 that

$$\int |(\phi, \pi(x)\psi)\,\mathrm{conj}(\phi', \pi'(x)\psi')|\, dx^* \leqq (d_\pi d_{\pi'})^{-1} |\phi|^2 |\psi|^2 |\phi'|^2 |\psi'|^2.$$

Therefore for any given $\phi \in \mathfrak{H}$ and $\phi' \in \mathfrak{H}'$, there exists a bounded linear operator A from $\mathfrak{H}$ to $\mathfrak{H}'$ such that

$$(\psi', A\psi) = \int (\phi, \pi(x)\psi)\,\mathrm{conj}(\phi', \pi'(x)\psi')\, dx^*$$

for all $\psi' \in \mathfrak{H}'$ and $\psi \in \mathfrak{H}$. It follows immediately from this relation that

$$(\psi', A\pi(x)\psi) = (\pi'(x^{-1})\psi', A\psi) \qquad (x \in G)$$

and therefore $A\pi(x) = \pi'(x)A$. In order to prove the first statement of the

[2] conj c denotes the conjugate of a complex number c.

theorem it would be enough to show that if $A \neq 0$, π and π' are equivalent. So let us suppose $A \neq 0$. Choose ψ in $\mathfrak{H}$ such that $A\psi \neq 0$. Then if $h \in C_c(G)$ it is clear that

$$A\pi(h)\psi = \pi'(h)A\psi$$

and therefore from the Corollary to Lemma 2, π and π' are infinitesimally equivalent. But since they are both unitary this implies that they are equivalent [5(b), Theorem 8].

In order to prove the second statement we may assume that $\pi' = \pi$ since

$$(\phi', \pi'(x)\psi') = (U^{-1}\phi', \pi(x)U^{-1}\psi')$$

in the general case. Hence if we keep to the above notation, A is now a bounded linear operator on $\mathfrak{H}$, which commutes with $\pi(x)$ $(x \in G)$. Since π is irreducible, A must be a scalar multiple of the identity. Hence

$$\int (\phi, \pi(x)\psi)\,\mathrm{conj}(\phi', \pi(x)\psi')\,dx^* = c_{\phi,\phi'}(\psi', \psi)$$

where $c_{\phi,\phi'}$ is a complex number depending only on ϕ and ϕ'. But obviously

$$\int (\phi, \pi(x)\psi)\,\mathrm{conj}(\phi', \pi(x)\psi')\,dx^* = \int (\psi', \pi(x)\phi')\,\mathrm{conj}(\psi, \pi(x)\phi)\,dx^*$$

and therefore

$$c_{\phi,\phi'}(\psi', \psi) = c_{\psi,\psi'}(\phi, \phi').$$

Choose $\psi' = \psi = \psi_0 \neq 0$ and put $c = (\psi_0, \psi_0)/|\psi_0|^2$. Then $c_{\phi,\phi'} = c(\phi', \phi)$ for all $\phi', \phi \in \mathfrak{H}$. In particular if we put $\phi = \phi' = \psi = \psi' = \psi_0$ we find that $c = d_\pi^{-1}$. Thus the theorem is proved.

4. The character of a square-integrable representation. Let $C_c^\infty(G)$ denote, as before, the subspace of $C_c(G)$ consisting of those functions which are indefinitely differentiable everywhere. For $x^* \in G^*$, we put $y^{x^*} = xyx^{-1}$ $(y \in G)$ where x is any element in G whose image in G^* is x^*.

THEOREM 2. *Let π be a square-integrable representation of G on a Hilbert space and let T_π denote the character* [5(c)] *of π. Then if $f \in C_c^\infty(G)$,*

$$T_\pi(f) = d_\pi \int_{G^*} dx^* \{ \int_G f(y^{x^*})\,(\phi, \pi(y)\phi)\,dy \}$$

where ϕ is any unit vector in $\mathfrak{H}$ and d_π is the formal degree of π.

Let Q denote the operator $\int_G f(y)\pi(y)\,dy$. We know [5(c), p. 243] that

there exists a complete orthonormal set $\{\psi_j\}_{j\in J}$ in $\mathfrak{H}$ such that $\sum_{i,j\in J} |Q_{ij}| < \infty$ where $Q_{ij} = (\psi_i, Q\psi_j)$. Moreover $\pi(x^{-1})Q\pi(x)$ $(x \in G)$ depends only on x^* and so we may denote it by Q^{x^*}. Then

$$(\phi, Q^{x^*}\phi) = (\pi(x)\phi, Q\pi(x)\phi) = \sum_i (\pi(x)\phi, \psi_i)(\psi_i, Q\pi(x)\phi)$$
$$= \sum_i \sum_j (\pi(x)\phi, \psi_i) Q_{ij} (\psi_j, \pi(x)\phi).$$

But if we make use of the Schwartz inequality and the Schur orthogonality relations, we get

$$\sum_{i,j} \int_{G^*} |(\pi(x)\phi, \psi_i) Q_{ij} (\psi_j, \pi(x)\phi)| \, dx^*$$
$$\leqq \sum_{i,j} |Q_{ij}| \{ \int_{G^*} |(\pi(x)\phi, \psi_i)|^2 \, dx^* \int_{G^*} |(\psi_j, \quad (x)\phi)|^2 \, dx^* \}^{\frac{1}{2}}$$
$$= d_\pi^{-1} \sum_{i,j} |Q_{ij}| < \infty$$

and therefore by Lebegue's Theorem the above series for $(\phi, Q^{x^*}\phi)$ may be integrated over G^* term by term. Hence

$$\int_{G^*} (\phi, Q^{x^*}\phi) \, dx^* = \sum_i \sum_j Q_{ij} \int_{G^*} (\pi(x)\phi, \psi_i)(\psi_j, \pi(x)\phi) \, dx^*$$
$$= \sum_i \sum_j (Q_{ij}/d_\pi)(\psi_j, \psi_i) = d_\pi^{-1} \sum Q_{ii}$$
$$= d_\pi^{-1} \operatorname{Sp} Q = d_\pi^{-1} T_\pi(f).$$

But

$$(\phi, Q^{x^*}\phi) = \int_G f(y^{x^*})(\phi, \pi(y)\phi) \, dy$$

and so the theorem is proved.

It should be noticed that, in general, the double integral in the above theorem is not absolutely convergent and therefore the order of the two integrations cannot be interchanged.

5. The discrete part of the Plancherel measure. We shall assume in this section that Z is finite and $Z_0 = \{1\}$. Let $\mathcal{E}$ denote the set of all equivalence classes of irreducible unitary representations of G. We consider the Hilbert space $L_2(G)$ consisting of all complex-valued functions on G which are square-integrable with respect to the Haar measure. Let λ denote the left regular representation of G on $L_2(G)$ defined by

$$(\lambda(x)f)(y) = f(x^{-1}y) \qquad (x, y \in G; f \in L_2(G)).$$

Then λ is unitary. We say that a class $\omega \in \mathcal{E}$ is discrete if there exists a closed subspace $\mathfrak{H} \neq 0$ of $L_2(G)$ which is invariant and irreducible under $\lambda(G)$ and such that the corresponding representation of G on $\mathfrak{H}$ lies in ω. It is known (see Godemont [4(a), Theorem 1]) that ω is discrete if and only if every representation in ω is square-integrable. Let $\mathcal{E}_0$ denote the set of all discrete classes in $\mathcal{E}$. If $\omega \in \mathcal{E}_0$, we denote by d_ω the formal degree of any representation in ω.

For any $\omega \in \mathcal{E}$, let T_ω denote the character [5(c)] of any representation in ω. Then it is known (see Segal [10(b)], Mautner [8] and [5(b), Theorem 7]) that there exists a unique positive measure μ on $\mathcal{E}$ such that

$$\int_G |f(x)|^2 \, dx = \int_{\mathcal{E}} T_\omega(\tilde{f} * f) \, d\mu \qquad (f \in C_c(G))$$

where $\tilde{f}(x) = \operatorname{conj} f(x^{-1})$ $(x \in G)$. We shall call μ the Plancherel measure on $\mathcal{E}$. First we prove the following simple lemma.

LEMMA 5. *Every single point ω_0 in $\mathcal{E}$ is μ-measurable.*

For any $f \in C_c(G)$ put $\|f\|_1 = \int_G |f(x)| \, dx$. Then under this norm $C_c(G)$ becomes a separable metric space. Let π_0 be a representation in ω_0 and let $\mathfrak{M}_{\omega_0}$ denotes the set of all $f \in C_c(G)$ such that $\pi_0(f) = 0$. Then $\mathfrak{M}_{\omega_0}$ is also separable under the above metric and so we can select a sequence $\{\alpha_1, \alpha_2, \cdots\}$ in $\mathfrak{M}_{\omega_0}$ which is dense in $\mathfrak{M}_{\omega_0}$. Put $F_n(\omega) = T_\omega(\tilde{\alpha}_n * \alpha_n)$ $(\omega \in \mathcal{E})$. Then $F_1, F_2, \cdots$ are all measurable functions on $\mathcal{E}$ and it follows from the Corollary to Lemma 2 that ω_0 is the only point in $\mathcal{E}$ where they vanish simultaneously. From this the lemma follows immediately.

Our main object in this section is to prove the following theorem.

THEOREM 3. *If ω_0 is a discrete class, $\mu(\omega_0) = d_{\omega_0}$.*

Define π_0, $\mathfrak{M}_{\omega_0}$ and $\{\alpha_n\}_{n \geqq 1}$ as above and let ψ_0 be a nonzero vector in the representation space $\mathfrak{H}$ of π_0. Put $g(x) = (\pi_0(x)\psi_0, \psi_0)$. Then since ω_0 is discrete, $g \in L_2(G)$. We first need the following lemma.

LEMMA 6. *There exists a sequence $g_1, g_2, \cdots$ in $C_c(G)$ and a subset $\mathcal{E}'$ of $\mathcal{E}$ satisfying the following conditions:*

(1) *The complement of $\mathcal{E}'$ in $\mathcal{E}$ is of μ-measure zero.*

(2) $\operatorname{Lim}_{n \to \infty} g_n = g$ *in* $L_2(G)$.

(3) *For every* $\omega \in \mathcal{E}'$, $\lim_{n\to\infty} T_\omega\ (\tilde{g}_n * g_n)$ *exists and is finite and*[3] $\lim_{n\to\infty} T_\omega((\alpha_m * g_n)^\sim * (\alpha * g_n)) = 0$ *for every* $m \geqq 1$.

Since $C_c(G)$ is dense in $L_2(G)$ we can choose a sequence $\{h_1, h_2, \cdots\}$ in $C_c(G)$ such that $h_n \to g$ in $L_2(G)$. Then

$$\int_{\mathcal{E}} T_\omega(\tilde{h}_n * h_n)\, d\mu = \| h_n \|^2 \to \| g \|^2$$

where $\| \ \|$ denotes the norm in $L_2(G)$. Therefore by the Riesz-Fischer Theorem, there exists a subset $\mathcal{E}_0' \subset \mathcal{E}$ and a subsequence $\{h_n^{(0)}\}$ of $\{h_n\}$ such that[4] (1) $\mathcal{E} - \mathcal{E}_0'$ is of μ-measure zero and (2) $\lim_{n\to\infty} T\,(h_n^{(0)\sim} * h_n^{(0)})$ exists and is finite for every $\omega \in \mathcal{E}_0'$. Now for each integer $r \geqq 0$ we shall define a subsequence $\{h_n^{(r)}\}$ of $\{h_n^{(0)}\}$ and a subset $\mathcal{E}_r' \subset \mathcal{E}_0'$ such that (1) $\mathcal{E} - \mathcal{E}_r'$ is of μ-measure zero and (2) $\lim_{n\to\infty} T_\omega((\alpha_m * h_n^{(r)})^\sim * (\alpha_m * h_n^{(r)})) = 0$ for $\omega \in \mathcal{E}_r'$ and $1 \leqq m \leqq r$. This has already been done for $r = 0$. So assuming that $\{h_n^{(r)}\}$ and $\mathcal{E}_r'$ have been defined, we proceed to define $\{h_n^{(r+1)}\}$ and $\mathcal{E}_{r+1}'$ by induction. Suppose $\omega \in \mathcal{E}_r'$. Then $T_\omega(h_n^{(0)\sim} * h_n^{(0)})$ converges to a finite limit and therefore if $\pi \in \omega$, $\pi(h_n^{(r)})$ converges with respect to the Hilbert-Schmidt (H. S.) norm to a bounded operator A_π. Now put $h_n' = \alpha_{r+1} * h_n^{(r)}$. Then $\pi(h_n') = \pi(\alpha_{r+1})\pi(h_n^{(r)})$ and since $\pi(\alpha_{r+1})$ is a bounded operator, it is obvious that $\pi(h_n')$ also converges to $\pi(\alpha_{r+1})A_\pi$ in the H. S. norm. On the other hand since $\| h_n^{(r)} - g \| \to 0$, $\alpha_{r+1} * h_n^{(r)} \to \alpha_{r+1} * g$ in $L_2(G)$. But

$$(\alpha_{r+1} * g)(x) = (\pi_0(x)\psi_0, \pi_0(\alpha_{r+1})\psi_0) = 0 \qquad (x \in G)$$

because $\alpha_{r+1} \in \mathfrak{M}_{\omega_0}$. Therefore $\alpha_{r+1} * h_n^{(r)} \to 0$ in $L_2(G)$ and so

$$\int_{\mathcal{E}} T_\omega((\alpha_{r+1} * h_n^{(r)})^\sim * (\alpha_{r+1} * h_n^{(r)}))\, d\mu \to 0$$

as $n \to \infty$. Hence by the Riesz-Fischer Theorem we can select a subsequence $\{h_n^{(r+1)}\}$ of $\{h_n^{(r)}\}$ and a subset $\mathcal{E}_{r+1} \subset \mathcal{E}$ such that (1) $\mathcal{E} - \mathcal{E}_{r+1}$ is of μ-measure zero and (2) if $\omega \in \mathcal{E}_{r+1}$,

$$\lim_{n\to\infty} T_\omega((\alpha_{r+1} * h_n^{(r+1)})^\sim * (\alpha_{r+1} * h_n^{(r+1)})) = 0.$$

Now put $\mathcal{E}_{r+1}' = \mathcal{E}_r' \cap \mathcal{E}_{r+1}$. Then $\mathcal{E}_{r+1}$ and $\{h_n^{(r+1)}\}$ satisfy all the required conditions. Hence if we take $\mathcal{E}' = \bigcap_{r=1}^{\infty} \mathcal{E}_r'$ and $g_n = h_n^{(n)}$ we get the assertion of the lemma.

---

[3] We write $f^\sim$ instead of $\tilde{f}$ $(f \in C_c(G))$ whenever it is convenient to do so.

[4] $\mathcal{E} - \mathcal{E}_0'$ is the complement of $\mathcal{E}_0'$ in $\mathcal{E}$.

Let us now come to the proof of the theorem. Let ω be a class in $\mathcal{E}'$ and π a representation in ω. Since $T_\omega(\tilde{g}_n * g_n)$ is convergent, the operators $\pi(g_n)$ converge to a limit in the H. S. norm. Let A_π denote this limit. Since $\pi(\alpha_r)$ is a bounded operator $\pi(\alpha_r * g_n) = \pi(\alpha_r)\pi(g_n)$ also converges to $\pi(\alpha_r)A_\pi$ in the H. S. norm. However

$$\operatorname*{Lim}_{n\to\infty} T_\omega((\alpha_r * g_n)^\sim * (\alpha_r * g_n)) = 0$$

and therefore $\pi(\alpha_r)A_\pi = 0$ $(r \geqq 1)$. Now suppose $\omega \neq \omega_0$. Then if ψ lies in the representation space of π, $\pi(\alpha_r)A_\pi\psi = 0$. Since π and π_0 are not equivalent (and therefore also not infinitesimally equivalent [5(b), Theorem 8]), it follows from the Corollary to Lemma 2 that $A_\pi\psi = 0$. This being true for every ψ, $A_\pi = 0$. But $\pi(g_n)$ tends to A_π in the H. S. norm and so this shows that

$$\operatorname*{Lim}_{n\to\infty} T_\omega(\tilde{g}_n * g_n) = 0.$$

On the other hand $g_n \to g$ in $L_2(G)$ and therefore

$$\| g \|^2 = \operatorname*{Lim}_{n\to\infty} \int_{\mathcal{E}} T_\omega(\tilde{g}_n * g_n)\, d\mu.$$

From this it follows that $\omega_0 \in \mathcal{E}'$. For otherwise, in view of what we have just said, $\operatorname*{Lim}_{n\to\infty} T_\omega(\tilde{g}_n * g_n) = 0$ for all $\omega \in \mathcal{E}'$ and since $\mathcal{E} - \mathcal{E}'$ is of μ-measure zero we would have

$$\| g \|^2 = \operatorname*{Lim}_{n\to\infty} \int_{\mathcal{E}} T_\omega(\tilde{g}_n * g_n)\, d\mu = 0.$$

But this is false since g is continuous and $g(1) = |\psi_0|^2 \neq 0$. Therefore $\omega_0 \in \mathcal{E}'$ and so $T_{\omega_0}(\tilde{g}_n * g_n)$ tends to a finite limit. This shows that

$$\begin{aligned} \| g \|^2 &= \operatorname*{Lim}_{n\to\infty} \int_{\mathcal{E}'} T_\omega(\tilde{g}_n * g_n)\, d\mu = \int_{\mathcal{E}'} \operatorname*{Lim}_{n\to\infty} T_\omega(\tilde{g}_n * g_n)\, d\mu \\ &= \mu(\omega_0) \operatorname*{Lim}_{n\to\infty} T_{\omega_0}(\tilde{g}_n * g_n). \end{aligned}$$

Now let[5] $\{\psi_1, \psi_2, \cdots\}$ be a complete orthonormal set in the representation space $\mathfrak{H}$ of π_0. Then if $A_n = \pi_0(g_n)$,

$$(\psi_i, A_n\psi_j) = \int_G g_n(x)\,(\psi_i, \pi_0(x)\psi_j)\, dx.$$

Since π_0 is square-integrable it follows that

$$\operatorname*{Lim}_{n\to\infty} (\psi_i, A_n\psi_j) = \int g(x)(\psi_i, \pi_0(x)\psi_j)dx = (\psi_i, \pi_0(g)\psi_j).$$

[5] It is not difficult to see that $\mathfrak{H}$ is separable.

But we have seen above that A_n converges in the H. S. norm and so its limit must be $\pi_0(g)$. This proves that

$$\lim_{n\to\infty} T_{\omega_0}(\tilde{g}_n * g_n) = T_{\omega_0}(\tilde{g} * g)$$

and therefore

$$\| g \|^2 = \mu(\omega_0) T_{\omega_0}(\tilde{g} * g).$$

But

$$(\psi_i, \pi_0(g)\psi_j) = \int (\pi_0(x)\psi_0, \psi_0)(\psi_i, \pi_0(x)\psi_j)\, dx$$

$$= d_{\omega_0}^{-1}(\psi_0, \psi_j)(\psi_i, \psi_0).$$

Hence

$$T_{\omega_0}(\tilde{g} * g) = \sum_{i,j} |(\psi_i, \pi_0(g)\psi_j)|^2 = d_{\omega_0}^{-2} \sum_{i,j} |(\psi_0, \psi_j)(\psi_i, \psi_0)|^2 = d_{\omega_0}^{-2} |\psi_0|^4.$$

On the other hand

$$\| g \|^2 = \int |(\pi(x)\psi_0, \psi_0)|^2\, dx = d_{\omega_0}^{-1} |\psi_0|^4$$

and therefore $\mu(\omega_0) = d_{\omega_0}$.

COROLLARY 1. *Let π be a square-integrable representation of G. Then if $f \in C_c(G)$,*

$$\| \int f(x)\pi(x)\, dx \|^2 \leqq d_\pi^{-1} \int |f(x)|^2\, dx$$

where $\| A \|$ denotes the H. S. *norm of an operator A.*

For $\int |f(x)|^2\, dx = \int_{\mathcal{E}} T_\omega(\tilde{f} * f)\, d\mu \geqq \mu(\omega_0) T_{\omega_0}(\tilde{f} * f)$ if ω_0 is the class of π. But $\mu(\omega_0) = d_\pi$ and

$$T_{\omega_0}(\tilde{f} * f) = \| \int f(x)\pi(x)\, dx \|^2.$$

Hence the result

COROLLARY 2. *A class $\omega_0 \in \mathcal{E}$ is discrete if and only if $\mu(\omega_0) > 0$.*

We have seen that if ω_0 is discrete $\mu(\omega_0) = d_{\omega_0}$ is positive. Conversely suppose $\mu(\omega_0)$ is positive. Let π be a representation in ω_0 and ϕ a unit vector in the representation space of π. Then if $f \in C_c(G)$,

$$|(\phi, \pi(f)\phi)|^2 \leqq \| \int f(x)\pi(x)\, dx \|^2 = T_{\omega_0}(f * f).$$

On the other hand

$$\| f \|^2 = \int_{\mathcal{E}} T_\omega(\tilde{f} * f)\, d\mu \geqq \mu(\omega_0) T_{\omega_0}(\tilde{f} * f).$$

Therefore

$$|(\phi, \pi(f)\phi)|^2 \leqq \mu(\omega_0)^{-1} \| f \|^2.$$

Let $L_1(G)$ denote, as usual, the space of all functions which are integrable on G. Then if $L = L_1(G) \cap L_2(G)$, it follows from the above inequality that

$$|(\phi, \pi(g)\phi)|^2 \leqq \mu(\omega_0)^{-1} \| g \|^2 \qquad (g \in L)$$

where $\pi(g) = \int g(x)\pi(x)dx$. U being any compact neighborhood of 1 in G, we now define a function $g_U \in L$ as follows. $g_U(x) = (\pi(x)\phi, \phi)$ if $x \in U$ and $g_U(x) = 0$ otherwise. Then

$$(\phi, \pi(g_U)\phi) = \int g_U(x)\,(\phi, \pi(x)\phi)\,dx = \int_U |(\phi, \pi(x)\phi)|^2\,dx = \| g_U \|^2$$

and therefore

$$\| g_U \|^2 = \int_U |(\phi, \pi(x)\phi)|^2\,dx \leqq \mu(\omega_0)^{-1}.$$

Hence

$$\int |(\phi, \pi(x)\phi)|^2\,dx = \sup_U \int_U |(\phi, \pi(x)\phi)|^2\,dx \leqq \mu(\omega_0)^{-1}.$$

This proves that π is square-integrable and therefore ω_0 is discrete.

Part II.

6. Some algebraic results. We shall now study in detail certain special representations which have been constructed in another paper [5(f)] and prove that, under suitable conditions, they are square-integrable or even integrable. Later (in Sections 9 and 10) we shall also obtain a formula for the formal degree of these representations.

Let $\mathfrak{h}_0$ be a maximal abelian subalgebra of $\mathfrak{k}_0$. In accordance with the assumption of [5(e), (f)] *we shall suppose that* $\mathfrak{h}_0$ *is also maximal abelian in* $\mathfrak{g}_0$. From now on we use the notation and the terminology of [5(e), § 3] without further comment. Then $\mathfrak{h}$ is a Cartan subalgebra of $\mathfrak{g}$. Suppose an order has been introduced once for all in the space $\mathfrak{F}_R$ of real linear functions on $\mathfrak{h}$ (see [5(e), § 2]) and P is the set of all positive roots of $\mathfrak{g}$ (with respect to $\mathfrak{h}$) in this order. *We shall further assume that every non-compact root is totally positive.* Since we are now interested primarily in unitary representations, this assumption is justified in view of Corollary 1 to Lemma 19 of [5(e)].

Let $\mathfrak{k}$ be a subalgebra of $\mathfrak{g}$. Suppose there exists a set Q of totally positive roots such that

$$\mathfrak{l} = \mathfrak{l} \cap \mathfrak{k} + \sum_{\gamma \in Q} (CX_\gamma + CX_{-\gamma})$$

and let β be the lowest root in Q. Then X_β, $X_{-\beta}$ and therefore also $H_\beta = [X_\beta, X_{-\beta}]$ are in $\mathfrak{l}$. Let $\mathfrak{l}_\beta$ denote the centralizer of $CH_\beta + CX_\beta + CX_{-\beta}$ in $\mathfrak{l}$. It is obvious that $\mathfrak{l}_\beta$ is invariant under θ and therefore

$$\mathfrak{l}_\beta = \mathfrak{l}_\beta \cap \mathfrak{k} + \mathfrak{l}_\beta \cap \mathfrak{p}.$$

LEMMA 7. *$C(X_\beta + X_{-\beta}) + \mathfrak{l}_\beta \cap \mathfrak{p}$ is exactly the set of all elements in $\mathfrak{l} \cap \mathfrak{p}$ which commute with $X_\beta + X_{-\beta}$.*

Let Q' be the set of all roots in Q other than β. Then if $X \in \mathfrak{l} \cap \mathfrak{p}$,

$$X = c_\beta' X_\beta + c_{-\beta}' X_{-\beta} + \sum_{\gamma \in Q'} (c_\gamma X + c_{-\gamma} X_{-\gamma})$$

where c_β', $c_{-\beta}'$, c_γ, $c_{-\gamma}$ are complex numbers. Now

$$\mathfrak{g} = \mathfrak{h} + \sum_{\delta \in P} (CX_\delta + CX_{-\delta})$$

where the sum is direct. Hence it is clear that the component of $[X, X_\beta + X_{-\beta}]$ in $\mathfrak{h}$ is $(c_\beta' - c_{-\beta}')H_\beta$. So if X commutes with $X_\beta + X_{-\beta}$, $c_\beta' = c_{-\beta}'$ and therefore

$$Y = \sum_{\gamma \in Q} (c_\gamma X_\gamma + c_{-\gamma} X_{-\gamma})$$

also commutes with $(X_\beta + X_{-\beta})$. In order to prove the lemma it is enough to show that $Y \in \mathfrak{l}_\beta$. Let us suppose then that this is false. Define $c_\delta = 0$ for any root δ for which neither δ nor $-\delta$ is in Q'. Then it is obvious that there exists roots δ such that (1) $c_\delta \neq 0$ and (2) $X_\delta \notin \mathfrak{l}_\beta$, for otherwise Y would lie in $\mathfrak{l}_\beta$. Let δ_0 be the highest such root. Since $[Y, X_\beta + X_{-\beta}] = 0$, it follows that $\delta_0 + \beta$ is not a root. However $X_{\delta_0} \notin \mathfrak{l}_\beta$ and so $\delta_0 - \beta$ must be a root. The coefficient of $X_{\delta_0 - \beta}$ in $[Y, X_\beta]$ (with respect to the above decomposition of $\mathfrak{g}$ as a direct sum) is then clearly different from zero. This means that $\delta_0 - \beta = \gamma + \beta$ where γ is some root with $c_\gamma \neq 0$. Hence $\gamma = \delta_0 - 2\beta$ is a root and $X_{\delta_0 - 2\beta} \in \mathfrak{l}$. Since δ_0 and β are both noncompact, $\alpha = \delta_0 - \beta$ is compact. Moreover β being totally positive, $\delta_0 = \beta + \alpha$ and $2\beta - \delta_0 = \beta - \alpha$ are also totally positive (Lemma 12 of [5(e)]) and $X_{\beta+\alpha} = X_{\delta_0}$, $X_{\beta-\alpha} = X_{2\beta-\delta_0}$ are both in $\mathfrak{l}$. Therefore $\beta + \alpha$ and $\beta - \alpha$ are in Q. This however is impossible since β is the lowest root in Q and so the lemma is proved.

Let Q_β be the set of all $\gamma \in Q$ such that $\gamma \neq \beta$ and neither $\gamma+\beta$ nor $\gamma-\beta$ is a root. Then it is obvious that

$$\mathfrak{l}_\beta = \mathfrak{l}_\beta \cap \mathfrak{k} + \sum_{\gamma \in Q_\beta} (CX_\gamma + CX_{-\gamma}).$$

Therefore $\mathfrak{l}_\beta$ satisfies the same condition as the one imposed above on $\mathfrak{l}$. Now we shall define a sequence $\mathfrak{g} = \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \mathfrak{g}_3 \supset \cdots$ of subalgebras of $\mathfrak{g}$ such that each $\mathfrak{g}_r$ satisfies this condition. The inductive definition is as follows. If $\mathfrak{g}_r \subset \mathfrak{k}$, $\mathfrak{g}_{r+1} = \mathfrak{g}_r$. Otherwise let β be the lowest totally positive root such that $X_\beta \in \mathfrak{g}_r$. Then $\mathfrak{g}_{r+1}$ is the centralizer of $CH_\beta + CX_\beta + CX_{-\beta}$ in $\mathfrak{g}_r$. It is obvious that $\dim \mathfrak{g}_{r+1} < \dim \mathfrak{g}_r$ unless $\mathfrak{g}_r \subset \mathfrak{k}$ and therefore $\mathfrak{g}_r \subset \mathfrak{k}$ if r is sufficiently large. Let $s \geqq 0$ be the least integer such that $\mathfrak{g}_{s+1} \subset \mathfrak{k}$ and let γ_r be the lowest totally positive root such that $X_{\gamma_r} \in \mathfrak{g}_r$ $(1 \leqq r \leqq s)$. One proves easily by induction on r that if $X_\alpha \in \mathfrak{g}_r$ for some root α then $X_{-\alpha}$ is also in $\mathfrak{g}_r$.

LEMMA 8. *$\gamma_i \pm \gamma_j$ $(1 \leqq i < j \leqq s)$ is never a root or zero and the elements $(X_{\gamma_i} + X_{-\gamma_i})$ $i = 1, 2, \cdots, s$ span a maximal abelian subspace of $\mathfrak{p}$ over C.*

If $i < j$, $\mathfrak{g}_{i+1} \supset \mathfrak{g}_j$ and therefore $\mathfrak{g}_j$ commutes with X_{γ_i} and $X_{-\gamma_i}$. Hence $\gamma_i \pm \gamma_j$ is not a root or zero. Let $\mathfrak{a}_\mathfrak{p}$ be the subspace of $\mathfrak{p}$ spanned by $(X_{\gamma_i} + X_{-\gamma_i})$ $i = 1, 2, \cdots, s$. Then $\mathfrak{a}_\mathfrak{p}$ is obviously abelian. Let X be an element in $\mathfrak{p}$ which commutes with $\mathfrak{a}_\mathfrak{p}$. We have to show that $X \in \mathfrak{a}_\mathfrak{p}$. Suppose this is false. Then it is obvious that $X \notin \mathfrak{k} + \mathfrak{a}_\mathfrak{p}$. Since $\mathfrak{g}_{s+1} \subset \mathfrak{k}$, we can choose r $(1 \leqq r \leqq s)$ such that $X \in \mathfrak{g}_r + \mathfrak{a}_\mathfrak{p}$ but $X \notin \mathfrak{g}_{r+1} + \mathfrak{a}_\mathfrak{p}$. Let $X = Y + Z$ $(Y \in \mathfrak{g}_r, Z \in \mathfrak{a}_\mathfrak{p})$. Since X commutes with $X_{\gamma_r} + X_{-\gamma_r}$, the same holds for Y. Also $Y = X - Z \in \mathfrak{g}_r \cap \mathfrak{p}$. Therefore we conclude from Lemma 7 that

$$Y = c(X_{\gamma_r} + X_{-\gamma_r}) + Y_1$$

where $Y_1 \in \mathfrak{g}_{r+1} \cap \mathfrak{p}$ and $c \in C$. Then $Z_1 = Z + c(X_{\gamma_r} + X_{-\gamma_r})$ lies in $\mathfrak{a}_\mathfrak{p}$ and so

$$X = Y_1 + Z_1 \in \mathfrak{g}_{r+1} + \mathfrak{a}_\mathfrak{p}.$$

Since this contradicts the definition of r, the lemma follows.

COROLLARY. *Let $\mathfrak{a}_{\mathfrak{p}_0} = \sum_{i=1}^{s} R(X_{\gamma_i} + X_{-\gamma_i})$. Then $\mathfrak{a}_{\mathfrak{p}_0} = \mathfrak{p}_0 \cap \mathfrak{a}_\mathfrak{p}$ and therefore it is a maximal abelian subspace of $\mathfrak{p}_0$.*

We know that $\tilde{\theta}(X_\gamma) = -X_{-\gamma}$ for any root γ (see [5(e), § 4]. There is a small mistake on p. 757 of [5(e)]. In line 22 $\tilde{\theta}$ should be replaced by η

which is the conjugation of $\mathfrak{g}$ with respect to $\mathfrak{g}_0$.). Hence $X_{\gamma_i}+X_{-\gamma_i}\in\mathfrak{p}_0$. Moreover if

$$X=\sum_{i=1}^{s}c(X_{\gamma_i}+X_{-\gamma_i})\in\mathfrak{p}_0 \qquad (c_i\in C),$$

$X=-\tilde{\theta}(X)$ and therefore $c_i\in R$.

We now need some simple facts about a three dimensional Lie algebra.

LEMMA 9. *Let* I *be the Lie algebra of dimension* 3 *spanned over* C *by the elements* H,X,Y *satisfying the following relations:*

$$[X,Y]=H, \qquad [H,X]=2X, \qquad [H,Y]=-2Y.$$

Let ν *denote the automorphism of* I *given by*

$$\nu(Z)=\exp\tfrac{\pi}{4}\operatorname{ad}(X-Y))Z \qquad (Z\in\mathrm{I}).$$

Then $\nu(H)=-(X+Y)$, $\nu(X+Y)=H$, $\nu(X-Y)=X-Y$. *Moreover if* L *is any complex analytic group with the Lie algebra* I,

$$\exp t(X+Y)=\exp(zY)\exp(\log(\cosh t)H)\exp zX \qquad (t\in C,\cosh t\neq 0)$$

where[6] $z=\tanh t$.

It is well known that I is isomorphic to the Lie algebra of the group of all 2×2 complex matrices with determinant 1. Since this group is simply connected, it is enough to prove the above relations in it. Therefore we may identify X, Y, H with matrices as follows:

$$X=\begin{pmatrix}0&1\\0&0\end{pmatrix},\quad Y=\begin{pmatrix}0&0\\1&0\end{pmatrix},\quad H=\begin{pmatrix}1&0\\0&-1\end{pmatrix}.$$

The required relations are now verified by a simple calculation.

Let P_+ be all the totally positive roots of $\mathfrak{g}$. Then $\gamma_i\in P_+$ $1\leqq i\leqq s$. Consider the automorphism ν of $\mathfrak{g}$ given by $\nu=\exp\frac{\pi}{4}\operatorname{ad}(\sum_{j=1}^{s}(X_{\gamma_i}-X_{-\gamma_i}))$. It follows from Lemmas 8 and 9 that $\nu(X_{\gamma_j}+X_{-\gamma_j})=H_{\gamma_j}$ and therefore $\nu(\mathfrak{a}_\mathfrak{p})=\sum_{i=1}^{s}CH_{\gamma_i}$. This shows that H_{γ_i} and therefore also γ_i $(1\leqq i\leqq s)$ are

---

[6] Our result is valid with any determination of the logarithm. But for the sake of definiteness let us make the following convention. Choose a fixed square root of -1 in C and denote it by $(-1)^{\frac{1}{2}}$. Then if z is a non-zero complex number

$$\log z=\log|z|+(-1)^{\frac{1}{2}}\phi$$

where $\log|z|$ and ϕ are real and $0\leqq\phi<2\pi$. Hence in particular if z is real and positive $\log z$ is real.

linearly independent. Let $\mathfrak{a}_\mathfrak{k}$ be the orthogonal complement of $\nu(\mathfrak{a}_\mathfrak{p})$ in $\mathfrak{h}$ with respect to the positive definition Hermitian form—$B(\tilde{\theta}(X), X)$ ($X \in \mathfrak{g}$, see [5(e), § 4]). Since $\tilde{\theta}(H_\gamma) = -H_\gamma$ for every root γ, it is obvious that $B(H_{\gamma_i}, H) = 0$ and therefore $\gamma_i(H) = 0$ $1 \leqq i \leqq s$ if $H \in \mathfrak{a}_\mathfrak{k}$. This means that $X_{\gamma_i} - X_{-\gamma_i}$ $1 \leqq i \leqq s$ commute with H and therefore $\nu(H) = H$ ($H \in \mathfrak{a}_\mathfrak{k}$). Hence if $\mathfrak{a} = \mathfrak{a}_\mathfrak{p} + \mathfrak{a}_\mathfrak{k}$, $\nu(\mathfrak{a}) = \nu(\mathfrak{a}_\mathfrak{p}) + \mathfrak{a}_\mathfrak{k} = \mathfrak{h}$. As ν is an automorphism, it follows that $\mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$.

Let α, β be two roots of $\mathfrak{g}$ and let k, k' be the largest nonnegative integers such that $\beta - k\alpha$ and $\beta + k'\alpha$ are roots. Then it is known (see Weyl [11(b)]) that $\beta(H_\alpha) = k - k'$ and $\beta + r\alpha$ is a root or zero for an integer r if and only if $-k \leqq r \leqq k'$ Moreover if s_α is the Weyl reflexion corresponding to α, $s_\alpha(\beta + k'\alpha) = \beta - k\alpha$. These facts should be constantly borne in mind during the following discussion.

LEMMA 10. *If* $\gamma, \delta \in P_+$, $\gamma(H_\delta) \geqq 0$.

For $\gamma + \delta$ is not a root (Lemma 11 of [5(e)]) and obviously it is not zero. Hence $\gamma(H_\delta) \geqq 0$.

LEMMA 11. *Let α be any root such that $H_\alpha \in \mathfrak{a}_\mathfrak{k}$. Then α is compact and $\gamma_i \pm \alpha$ $(1 \leqq i \leqq s)$ can never be a root.*

Without loss of generality we may assume that $\alpha > 0$. If α is not compact, it must be totally positive and therefore $\alpha + \gamma_i$ is not a root [5(e), Lemma 11]. Since $H_\alpha \in \mathfrak{a}_\mathfrak{k}$, $\gamma_i(H_\alpha) = 0$ and so it follows that $\gamma_i - \alpha$ also cannot be a root or zero. Hence X_α commutes with X_{γ_i}, $X_{-\gamma_i}$ $1 \leqq i \leqq s$. This however is impossible since $\mathfrak{a}_\mathfrak{p}$ is maximal abelian in $\mathfrak{p}$. So α must be compact.

Now consider the sequence $\mathfrak{g} = \mathfrak{g}_1 \supset \mathfrak{g}_2 \supset \cdots$ introduced above. We shall prove that $X_\alpha \in \mathfrak{g}_r$ for every r. For otherwise choose the least $r \geqq 0$ such that $X_\alpha \notin \mathfrak{g}_{r+1}$. Since $\mathfrak{g}_1 = \mathfrak{g}$, $r \geqq 1$ and $X_\alpha \in \mathfrak{g}_r$. In view of the fact that $X_\alpha \notin \mathfrak{g}_{r+1}$, it is clear that either $\gamma_r + \alpha$ or $\gamma_r - \alpha$ is a root. But γ_r is the lowest totally positive root γ such that $X_\gamma \in \mathfrak{g}_r$. Since $X_\alpha \in \mathfrak{g}_r$, $X_{-\alpha}$ also lies in $\mathfrak{g}_r$ and so if $\gamma_r - \alpha$ were a root, $X_{\gamma_r - \alpha}$ would also lie in $\mathfrak{g}_r$. Since α is compact and positive, $\gamma_r - \alpha$ is also totally positive and $\gamma_r - \alpha < \gamma_r$. As this contradicts the definition of γ_r, we conclude that $\gamma_r - \alpha$ is not a root and therefore $\gamma_r + \alpha$ is a root. But this implies that $\gamma_r(H_\alpha) < 0$ which, in its turn, contradicts the fact $\gamma_r(H) = 0$ for all $H \in \mathfrak{a}_\mathfrak{k}$. Hence the lemma.

LEMMA 12. *Let α be a root. Then for any i $(1 \leqq i \leqq s)$, $\gamma_i + \alpha$ and $\gamma_i - \alpha$ cannot both be roots.*

We may assume $\alpha > 0$. If α is noncompact $\gamma_i + \alpha$ cannot be a root [5(e), Lemma 11]. So we may assume that α is compact. Suppose then that for some i, $\gamma_i \pm \alpha$ are both roots. Then they are both totally positive and hence from Lemma 10,

$$\gamma_i(H_\delta) \pm \alpha(H_\delta) \geqq 0$$

for every $\delta \in P_+$ $(j \neq i)$. Then it follows from Lemma 8 that $\gamma_i(H_{\gamma_j}) = 0$ and therefore $\pm \alpha(H_{\gamma_j}) \geqq 0$. This means that $\alpha(H_{\gamma_j}) = 0$. On the other hand $\gamma_i + \alpha$, γ_i, $\gamma_i - \alpha$ are all roots and therefore it follows from Lemma 15 of [5(e)] that $\gamma_i + 2\alpha$ and $\gamma_i - 2\alpha$ are not roots. Hence $\gamma_i(H_\alpha) = 0$. But this implies that $\alpha(H_{\gamma_i}) = 0$ and therefore $\alpha(H_{\gamma_j}) = 0$ $1 \leqq j \leqq s$. This however means that $H_\alpha \in \mathfrak{a}_\mathfrak{l}$ and so we get a contradiction with Lemma 11.

Let λ and μ be two linear functions on $\mathfrak{h}$. We write $\lambda \sim \mu$ if $\lambda - \mu$ vanishes identically on $\nu(\mathfrak{a}_\mathfrak{v}) = \sum_{1 \leqq i \leqq s} CH_{\gamma_i}$.

LEMMA 13. *Let α be a positive compact root. Then there are only the following three mutually exclusive possibilities:*

(1) *$H_\alpha \in \mathfrak{a}_\mathfrak{l}$ and therefore $\alpha \sim 0$ and $\gamma_i \pm \alpha$ $(1 \leqq i \leqq s)$ is never a root.*

(2) *There exists a unique index i $(1 \leqq i \leqq s)$ such that $\alpha + \frac{1}{2}\gamma_i \sim 0$.*

(3) *There exists two unique indices i, j $(1 \leqq i < j \leqq s)$ such that $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$.*

Since the first case is covered by Lemma 11, we may assume that $H_\alpha \notin \mathfrak{a}_\mathfrak{l}$. Then $\alpha(H_{\gamma_i}) \neq 0$ for some i and therefore $X_\alpha \notin \mathfrak{g}_{s+1}$. Let i be the least index $(1 \leqq i \leqq s)$ such that $X_\alpha \notin \mathfrak{g}_i$. Since α is positive, $\gamma_i + \alpha$ is a root while $\gamma_i - \alpha$ is not (see the proof of Lemma 11). Now suppose $\gamma_j + \epsilon\alpha$ is a root for some j $(1 \leqq j \leqq s, \epsilon = \pm 1)$. If $j \neq i$ we claim $\epsilon = -1$. For otherwise suppose $\gamma_j + \alpha$ is a root. Then it follows from Lemmas 9 and 10 that $\alpha(H_{\gamma_i}) = \gamma_j(H_{\gamma_i}) + \alpha(H_{\gamma_i}) \geqq 0$. On the other hand since $\gamma_i + \alpha$ is a root while $\gamma_i - \alpha$ is not, it is clear that $\gamma_i(H_\alpha) < 0$ and therefore $\alpha(H_{\gamma_i}) < 0$. As this conflicts with our conclusion above, $\epsilon = -1$. So we have two cases. Either (1) $\gamma_j \pm \alpha$ is never a root for $j \neq i$ or (2) $\gamma_j - \alpha$ is a root for some $j \neq i$.

In the first case $\alpha(H_{\gamma_j}) = 0$ for all $j \neq i$. Moreover α, $\alpha + \gamma_i$ are roots while $\alpha - \gamma_i$ is not. Since γ_i and $\alpha + \gamma_i$ are both totally positive $\alpha + 2\gamma_i$ is not a root [5(e), Lemma 11]. Hence $\alpha(H_{\gamma_i}) = -1$. This shows that $\alpha(H_{\gamma_j}) + \frac{1}{2}\gamma_i(H_{\gamma_j}) = 0$ for all j $(1 \leqq j \leqq s)$ and therefore $\alpha + \frac{1}{2}\gamma_i \sim 0$.

Now consider the second case. Let j be the least index such that $\gamma_j - \alpha$ is a root. Then $j \neq i$ and in view of our definition of i, $j > i$. If k

is any index $(1 \leqq k \leqq s)$ other than i, j we claim $\gamma_k \pm \alpha$ cannot be a root. We have already seen this for $\gamma_k + \alpha$. So suppose $\gamma_k - \alpha$ is a root. Then $\gamma_k(H_{\gamma_j}) - \alpha(H_{\gamma_j}) \geqq 0$ (Lemma 10) and therefore $-\alpha(H_{\gamma_j}) \geqq 0$ (Lemma 8). On the other hand α, $\alpha - \gamma_j = -(\gamma_j - \alpha)$ are roots while $\alpha + \gamma_j$ is not. Therefore $\alpha(H_{\gamma_j}) > 0$ giving a contradiction. This proves that $\gamma_k \pm \alpha$ are never roots $(k \neq i, j)$. Moreover as we have seen above, α, $\alpha + \gamma_i$ are roots while $\alpha - \gamma_i$ and $\alpha + 2\gamma_i$ are not and therefore $\alpha(H_{\gamma_i}) = -1$. Similarly since α, $\alpha - \gamma_j$ are roots while $\alpha + \gamma_j$, $\alpha + 2\gamma_j$ are not, $\alpha(H_{\gamma_j}) = 1$. Finally $\alpha(H_{\gamma_k}) = 0$ $(k \neq i, j)$ in view of our result above. Hence it is clear that $\alpha(H_{\gamma_k}) = \frac{1}{2}\gamma_j(H_{\gamma_k}) - \frac{1}{2}\gamma_i(H_{\gamma_k})$ $(1 \leqq k \leqq s)$ and therefore $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$. Moreover since γ_k $(1 \leqq k \leqq s)$ vanish identically on $\mathfrak{a}_\mathfrak{k}$, it is clear that their restriction on $\nu(\mathfrak{a}_\mathfrak{p})$ are linearly independent. The uniqueness of the indices i and j in the second and third cases of our lemma and the mutual exclusiveness of the three possibilities are therefore obvious.

For any index i $(1 \leqq i \leqq s)$ let C_i denote the set of all compact roots α such that $\alpha + \frac{1}{2}\gamma_i \sim 0$. Similarly let P_i denote the set of all totally positive roots γ for which $\gamma \sim \frac{1}{2}\gamma_i$. If $\alpha \in C_i$, it follows from Lemma 13 that $-\alpha$ cannot be positive. This shows that C_i consists of positive roots.

LEMMA 14. *$\alpha \to \gamma_i + \alpha$ $(\alpha \in C_i)$ is a one-one mapping of C_i onto P_i $(1 \leqq i \leqq s)$.*

For any root β let s_β denote the Weyl reflexion corresponding to β. Now if $\alpha \in C_i$, $\alpha \sim -\frac{1}{2}\gamma_i$ and therefore $\alpha(H_{\gamma_i}) = -1$. Hence $s_{\gamma_i}\alpha = \alpha - \alpha(H_{\gamma_i})\gamma_i = \alpha + \gamma_i$ and so $\gamma_i + \alpha$ is a root which is obviously in P_i. Conversely if $\gamma \in P_i$, $\gamma \sim \frac{1}{2}\gamma_i$ and therefore $\gamma(H_{\gamma_i}) = 1$. Hence $s_{\gamma_i}\gamma = \gamma - \gamma_i$ is a root. But as γ and γ_i are both noncompact, $\alpha = \gamma - \gamma_i$ must be compact. Moreover $\alpha + \frac{1}{2}\gamma_i = \gamma - \frac{1}{2}\gamma_i \sim 0$ and therefore $\alpha \in C_i$. Since it is obvious from its definition that the mapping is one-one, the lemma is proved.

For any given pair of indices i, j $(1 \leqq i < j \leqq s)$, let C_{ij} denote the set of all compact roots α such that $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$. Similarly let P_{ij} denote the set of all $\gamma \in P_+$ such that $\gamma \sim \frac{1}{2}(\gamma_j + \gamma_i)$. Again we conclude from Lemma 13 that every root in C_{ij} is positive.

LEMMA 15. *$\alpha \to \gamma_i + \alpha$ $(\alpha \in C_{ij})$ is a one-one mapping of C_{ij} onto P_{ij}.*

Let $\alpha \in C_{ij}$. Then $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$ and therefore $\alpha(H_{\gamma_i}) = -1$. Hence $s_{\gamma_i}\alpha = \alpha + \gamma_i$ and so it is clear that $\gamma_i + \alpha \in P_{ij}$. Conversely if $\gamma \in P_{ij}$, $\gamma(H_{\gamma_i}) = 1$ and therefore $s_{\gamma_i}\gamma = \gamma - \gamma_i = \alpha$ (say). Then α is compact and $\alpha \sim \frac{1}{2}(\gamma_j - \gamma_i)$.

Let C_0 be the set of all positive roots α such that $\alpha \sim 0$. We know (Lemma 11) that every root in C_0 is compact.

LEMMA 16. *Let P_0 denote the set $(\gamma_1, \gamma_2, \cdots, \gamma_s)$. Then P is the disjoint union of C_0, C_i, C_{ij}, P_0, P_i, P_{ij} $(1 < i < j < s)$.*

Since $\gamma_1, \cdots, \gamma_s$ are linearly independent on $\nu(\mathfrak{a}_\mathfrak{p})$, it is obvious that these sets are all disjoint. Let Q be their union. Then if γ is any positive root, we have to show that $\gamma \in Q$. If γ is compact, this follows from Lemma 13. So now suppose $\gamma \in P_+$. Since $\mathfrak{g}_{s+1} \subset \mathfrak{k}$, we can choose an index i $(1 \leqq i \leqq s)$ such that $X_\gamma \in \mathfrak{g}_i$ but $X_\gamma \notin \mathfrak{g}_{i+1}$. Moreover since $\gamma_i \in P_0 \subset Q$, we may assume that $\gamma \neq \gamma_i$. Then it follows from the definition of γ_i that $\gamma > \gamma_i$. As both γ and γ_i are in P_+, $\gamma + \gamma_i$ is not a root [5(e), Lemma 11]. Therefore since $X_\gamma \notin \mathfrak{g}_{i+1}$, $\alpha = \gamma - \gamma_i$ must be a root which is then obviously compact and positive. Therefore we can apply Lemma 13 to α. Since $\gamma = \gamma_i + \alpha$ is a root, $\alpha \notin C_0$ (Lemma 11). Hence either $\alpha \sim -\frac{1}{2}\gamma_j$ or $\alpha \sim \frac{1}{2}(\gamma_k - \gamma_j)$ for some j or (k, j) $(1 \leqq j < k \leqq s)$. In the first case $\gamma \sim \gamma_i - \frac{1}{2}\gamma_j$ and so[7] $\gamma(H_{\gamma_j}) = 2\delta_{ij} - 1$. But we know from Lemma 10, that $\gamma(H_{\gamma_j}) \geqq 0$. Therefore $i = j$, $\gamma \sim \frac{1}{2}\gamma_i$ and $\gamma \in P_i$. In the second case $\gamma \sim \gamma_i + \frac{1}{2}(\gamma_k - \gamma_j)$ and $\gamma(H_{\gamma_j}) = \delta_{ij} - 1$ since $k \neq j$. Therefore again in view of the fact that $\gamma(H_{\gamma_j}) \geqq 0$, we conclude that $i = j$ and hence $\gamma \in P_{jk}$. This shows that $\gamma \in Q$ and therefore $Q = P$.

LEMMA 17. *Let α, β be two roots such that $\alpha = \frac{1}{2}(\gamma_j - \gamma_i)$ and $\beta = \frac{1}{2}(\gamma_k - \gamma_j)$ $(1 \leqq i, j, k \leqq s)$. Then they are both compact if $k \neq i$, $\beta + \alpha = \frac{1}{2}(\gamma_k - \gamma_i)$ is a root.*

Consider the scalar product $\langle \alpha, \beta \rangle$ (see [5(e), § 2]). Since $k \neq i$, it follows from Lemma 8 that $\langle \alpha, \beta \rangle = -\frac{1}{4}\langle \gamma_j, \gamma_j \rangle < 0$. Hence $\alpha(H_\beta) < 0$ and therefore $\alpha + \beta$ is a root. The compactness of α and β is an immediate consequence of Lemma 16.

Let us say that $\gamma_i \prec \gamma_j$ $(1 \leqq i, j \leqq s)$ if $\frac{1}{2}(\gamma_j - \gamma_i)$ is a positive root. The above lemma shows that this relation is transitive and therefore it defines a partial order in the set P_0. It is obvious from the definition of γ_i that $\gamma_i < \gamma_j$ if $i < j$. Hence $\gamma_i \prec \gamma_j$ implies $i < j$.

Let r_i and r_{ij} be the number of roots in C_i and C_{ij} $(1 \leqq i < j \leqq s)$ respectively. Then it follows from Lemmas 14 and 15 that these are also the number of roots in P_i and P_{ij} respectively. Put $2\rho_+ = \sum_{\beta \in P_+} \beta$. Then we have the following result.

---

[7] $\delta_{ij} = 1$ or 0 according as $i = j$ or not.

LEMMA 18. $2\rho_+(H_{\gamma_i}) = 2 + r_i + \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji}$ $(1\leqq i\leqq s)$.

Let Q_i be the union of P_i, P_{ij} $(i<j\leqq s)$ and P_{ji} $(1\leqq j<i)$. Put $2\rho_i = \sum_{\gamma\in Q_i}\gamma$. Then since $\gamma_j(H_{\gamma_i}) = 2\delta_{ji}$, it is obvious that

$$2\rho_i(H_{\gamma_i}) = r_i + \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji}.$$

Also if γ is a totally positive root which does not lie in Q_i, it follows from Lemma 16 that $\gamma(H_{\gamma_i}) = 0$ unless $\gamma = \gamma_i$. Therefore since $\gamma_i(H_{\gamma_i}) = 2$,

$$2\rho_i(H_{\gamma_i}) = 2 + 2\rho_i(H_{\gamma_i})$$

and this gives the result.

LEMMA 19. *r_{ij} $(1\leqq i<j\leqq s)$ is even if and only if $\frac{1}{2}(\gamma_j-\gamma_i)$ is not a root.*

Let θ' denote the automorphism $\nu\theta\nu^{-1}$ of $\mathfrak{g}$. Since $\theta(\mathfrak{a}) = \mathfrak{a}$ and $\mathfrak{h} = \nu(\mathfrak{a})$, it follows that $\theta'(\mathfrak{h}) = \mathfrak{h}$. Therefore if α is a root the linear function $H\to\alpha(\theta' H)$ $(H\in\mathfrak{h})$ is also a root. We denote it by $\theta'\alpha$. It is clear that $\theta' H = -H$ if $H\in\nu(\mathfrak{a}_\mathfrak{p})$ and $\theta' H = H$ if $H\in\mathfrak{a}_\mathfrak{k}$. Hence $\alpha\neq -\theta'\alpha$ unless α vanishes identically on $\mathfrak{a}_\mathfrak{k}$. Now suppose $\alpha\in C_{ij}$ $(1\leqq i<j\leqq s)$. Then $\alpha\sim\frac{1}{2}(\gamma_j-\gamma_i)$ and therefore it is obvious that $\theta'\alpha\sim-\frac{1}{2}(\gamma_j-\gamma_i)$. In view of Lemma 16, this implies that $-\theta'\alpha\in C_{ij}$. Hence the mapping $\alpha\to-\theta'\alpha$ defines a permutation of order 2 in the set C_{ij}. Moreover $\alpha\neq-\theta'\alpha$ unless $\alpha = \frac{1}{2}(\gamma_j-\gamma_i)$. Therefore if we pair off α and $-\theta'\alpha$ together, it follows immediately that r_{ij} is odd or even according as $\frac{1}{2}(\gamma_j-\gamma_i)$ is a root or not.

7. Digression on a theorem of Cartan. Put

$$\mathfrak{p}_+ = \sum_{\beta\in P_+} CX_\beta \text{ and } \mathfrak{p}_- = \sum_{\beta\in P_+} CX_{-\beta}.$$

Then $\mathfrak{p}_+$, $\mathfrak{p}_-$ are abelian subalgebras of $\mathfrak{g}$ [5(e), Lemma 11] and $\mathfrak{g}$ is the direct sum of $\mathfrak{k}$, $\mathfrak{p}_+$ and $\mathfrak{p}_-$. Let G_c denote the simply connected complex Lie group with the Lie algebra $\mathfrak{g}$ and let $\mathfrak{P}_c^+$, K_c, $\mathfrak{P}_c^-$ be its analytic subgroups corresponding to $\mathfrak{p}_+$, $\mathfrak{k}$, $\mathfrak{p}_-$ respectively. Also let G_0, K_0 be the real analytic subgroups of G_c corresponding to $\mathfrak{g}_0$, $\mathfrak{k}_0$ respectively. Then $(q,k,p)\to qkp$ $(q\in\mathfrak{P}_c^-, k\in K_c, p\in\mathfrak{P}_c^+)$ is a one-one regular holomorphic mapping of the complex manifold $\mathfrak{P}_c^-\times K_c\times\mathfrak{P}_c^+$ into G_c and G_0 is contained in $\mathfrak{P}_c^- K_c\mathfrak{P}_c^+$ [5(f), Lemmas 4 and 5].

LEMMA 20. *Let* $X = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t\in C)$. *Then*

$$\exp X = \exp Y \exp H \exp Z$$

in G_c *where*[6]

$$Y=\sum_{i=1}^{s}(\tanh t_i)X_{-i},\qquad Z=\sum_{i=1}^{s}(\tanh t_i)X_i\qquad H=\sum_{i=1}^{s}\log(\cosh t_i)H_i$$

provided $\cosh t_i\neq 0$ $1\leqq i\leqq s$.

This follows immediately from Lemmas 8 and 9.

Now if we put $(X,Y)=-B(\tilde{\theta}(X),Y)$ and $\|X\|=(X,X)^{\frac{1}{2}}$ $(X,Y\in\mathfrak{g})$, $\mathfrak{g}$ becomes a finite-dimensional Hilbert space. Moreover since $\mathrm{ad}X$ is nilpotent for $X\in\mathfrak{p}_-$, it is easy to see that $X\to\exp X$ $(X\in\mathfrak{p}_-)$ is a one-one regular holomorphic mapping of $\mathfrak{p}_-$ onto $\mathfrak{P}_c^-$. Let $q\to\log q$ $(q\in\mathfrak{P}_c^-)$ denote its inverse. For $x\in G_0$, let $\zeta(x)$ denote the unique element in $\mathfrak{P}_c^-$ such that $x\in\zeta(x)K_c\mathfrak{P}_c^+$. Then we have the following result.

LEMMA 21. $\|\log\zeta(x)\|$ *remains bounded as* x *varies in* G_0.

Let $\mathfrak{P}_0$ be the set of all elements in G_0 of the form $\exp X$ $(X\in\mathfrak{p}_0)$. Then it is known that $G_0=K_0\mathfrak{P}_0$ (see Cartan [2(b), p. 17], also Mostow [9]). Let $z\to\mathrm{Ad}(z)$ $(z\in G_c)$ denote the adjoint representation of G_c. It follows from the definition of $\tilde{\theta}$ that if $k\in K_0$, $\mathrm{Ad}(k)$ is a unitary operator on $\mathfrak{g}$. Now $\mathfrak{a}_{\mathfrak{p}_0}$ is a maximal abelian subspace of $\mathfrak{p}_0$ (Lemma 8) and therefore $\mathfrak{p}_0=\bigcup_{k\in K}\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ (see Lemma 33 and also Cartan [2(a), p. 359]). Hence $G_0=K_0\mathfrak{A}K_0$ where $\mathfrak{A}$ is the analytic subgroup of G_0 corresponding to $\mathfrak{a}_{\mathfrak{p}_0}$. Moreover since $[\mathfrak{k},\mathfrak{p}_-]\subset\mathfrak{p}_-$, it is obvious that $\zeta(kxk')=k\zeta(x)k^{-1}$ $(k,k'\in K_0,$ $x\in G_0)$. Therefore if $x=kak'$ $(k,k\in K_0;a\in\mathfrak{A})$,

$$\log\zeta(x)=\mathrm{Ad}(k)(\log\zeta(a))$$

and so

$$\|\log\zeta(x)\|=\|\log\zeta(a)\|.$$

Now suppose $a=\exp X$ where $X=\sum_{i=1}^{s}t_i(X_{\gamma_i}+X_{-\gamma_i})$ $(t_i\in R)$. Then from Lemma 26,

$$\log\zeta(a)=\sum_{i=1}^{s}(\tanh t_i)X_{-\gamma_i}$$

and therefore

$$\|\log\zeta(a)\|\leqq\sum_{i=1}^{s}\|X_{-\gamma_i}\|$$

since $|\tanh t|\leqq 1$ for real t. Thus

$$\|\log\zeta(x)\|\leqq\sum_{i=1}^{s}\|X_{-\gamma_i}\|$$

for all $x\in G_0$ and so the lemma is proved.

This result has the following significance in relation to the theory of bounded symmetric homogeneous domains of E. Cartan [2(c)]. We know that $G_0K_c\mathfrak{P}_c^+$ is open in $\mathfrak{P}_c^-K_c\mathfrak{P}_c^+$ and $G_0 \cap (K_c\mathfrak{P}_{c+}) = K_0$ (see [5(f), § 2]). Since $K_c\mathfrak{P}_c^+$ is a group and $\mathfrak{P}_c^- \cap (K_c\mathfrak{P}_c^+) = \{1\}$, we can identify $\mathfrak{P}_c^-$ with the factor space $(\mathfrak{P}_c^-K_c\mathfrak{P}_c^+)/K_c\mathfrak{P}_c^+$. In this way $G_0/K_0 = (G_0K_c\mathfrak{P}_c^+)/K_c\mathfrak{P}_c^+$ becomes an open submanifold of $\mathfrak{P}_c^-$. The above lemma then shows that this submanifold is equivalent to a *bounded* domain in the complex Euclidean space $\mathfrak{p}_-$. This fact had previously been verified by Cartan [2(c)] by using the classification of all real simple groups and constructing the domain in each case separately.

8. Transformation of certain integrals. Let $(-1)^{\frac{1}{2}}$ denote a fixed square-root of -1 in C and put $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}}\mathfrak{p}_0$. Then $\mathfrak{u}$ is a compact real form of $\mathfrak{g}$ (see [5(b), p. 187]). Let $\mathfrak{a}_{\mathfrak{p}_0}$ denote any (real) maximal abelian subspace of $\mathfrak{p}_0$. We denote by $\mathfrak{a}_\mathfrak{p}$ the complexification of $\mathfrak{a}_{\mathfrak{p}_0}$ in $\mathfrak{p}$. Define G_c, G_0, K_0 as in Section 7 and let U, $\mathfrak{A}$, $\mathfrak{A}^*$ and $\mathfrak{A}_c$ be the (real) analytic subgroups of G_c corresponding to $\mathfrak{u}$, $\mathfrak{a}_{\mathfrak{p}_0}$, $(-1)^{\frac{1}{2}}\mathfrak{a}_{\mathfrak{p}_0}$ and $\mathfrak{a}_\mathfrak{p}$ respectively. Then K_0, U and $\mathfrak{A}^*$ are compact (see § 12 and [5(f), § 2]). Put $\mathfrak{q} = [\mathfrak{g}, \mathfrak{a}_\mathfrak{p}]$. It is obvious that $\mathrm{Ad}(a)\mathfrak{q} = \mathfrak{q}$ for $a \in \mathfrak{A}_c$. We put

$$D(a) = \det(\mathrm{Ad}(a) - \mathrm{Ad}(a^{-1}))_\mathfrak{q} \qquad (a \in \mathfrak{A}_c)$$

where $(\mathrm{Ad}(a) - \mathrm{Ad}(a^{-1}))_\mathfrak{q}$ is the restriction of $\mathrm{Ad}(a) - \mathrm{Ad}(a^{-1})$ on $\mathfrak{q}$. Let dx, dk, du, da, da^* denote the Haar measures on G_0, K_0 U, and $\mathfrak{A}^*$ respectively. We assume that

$$\int_{K_0} dk = \int_U du = 1.$$

On the other hand da and da^* are normalized as follows. The metric on $\mathfrak{g}$ (see Section 7) defines a Euclidean metric on the real vector space $\mathfrak{a}_{\mathfrak{p}_0}$ which is given by $\| H \|^2 = B(H, H)$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$. Let dH denote the element of volume in $\mathfrak{a}_{\mathfrak{p}_0}$ corresponding to this Euclidean metric and put $e(H) = \exp(-1)^{\frac{1}{2}}H$. The mappings $H \to \exp H$ and $H \to e(H)$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$ define homomorphisms of the additive group $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$ and $\mathfrak{A}^*$ respectively and it is clear that these homomorphisms are local isomorphisms. Hence we can normalize the Haar measures da and da^* in such a way that $da = dH = da^*$ $(a = \exp H$ and $a^* = e(H)$, $H \in \mathfrak{a}_{\mathfrak{p}_0})$.

Let $\mathfrak{A}'$ and $\mathfrak{A}^{*\prime}$ be the sets of those points a in $\mathfrak{A}$ and $\mathfrak{A}^*$ respectively where $D(a) \neq 0$. Then both $\mathfrak{A}'$ and $\mathfrak{A}^{*\prime}$ have only a finite number of connected components (see Section 12). Let w and w^* respectively denote

their number. Moreover let $C_c(G_0)$ be the set of all continuous functions on G_0 which vanish outside a compact set.

LEMMA 22. *Let g be a continuous function on U and B_0^* a connected component of $\mathfrak{A}^{*\prime}$. Then*

$$\int_U g(u)du \int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^* = w^* \int_{B_0^*} |D(a^*)|^{\frac{1}{2}} da^* \int_{K_0 \times K_0} g(ka^*k')\,dkdk'.$$

Moreover we can normalize the Haar measure dx on G_0 in such a way that

$$\int_{G_0} f(x)dx = w \int_{B_0} |D(a)|^{\frac{1}{2}} da \int_{K_0 \times K_0} f(kak')\,dkdk'$$

for all $f \in C_c(G_0)$ and every connected component B_0 of $\mathfrak{A}'$. This normalization of dx and the numbers w, w^* and $\int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^*$ are independent of the choice of $\mathfrak{a}_{\mathfrak{p}_0}$.

Although the proof of this lemma is not difficult, due to some technical complications, it is rather long. Hence in order not to interrupt our main argument, we postpone it until Section 12.

Now we assume that $\mathfrak{a}_{\mathfrak{p}_0}$, $\mathfrak{a}_{\mathfrak{p}}$, $\mathfrak{a}_{\mathfrak{k}}$ and $\mathfrak{a}$ are defined as in Section 6 so that $\nu(\mathfrak{a}) = \mathfrak{h}$. Let Σ be the set of all roots of $\mathfrak{g}$ with respect to $\mathfrak{a}$, which do not vanish identically on $\mathfrak{a}_{\mathfrak{p}}$. Then it is obvious that

$$|D(\exp H)| = \Big|\prod_{\alpha \in \Sigma} (e^{\alpha(H)} - e^{-\alpha(H)})\Big| \qquad (H \in \mathfrak{a}_{\mathfrak{p}}).$$

Now every linear function λ on $\mathfrak{h}$ defines a linear function λ' on $\mathfrak{a}$ by the rule $\lambda'(H) = \lambda(\nu(H))$ $(H \in \mathfrak{a})$. Moreover since $\mathfrak{a} \cap \mathfrak{h} = \mathfrak{a}_{\mathfrak{k}}$ and $\nu(H) = H$ for $H \in \mathfrak{a}_{\mathfrak{k}}$, λ and λ' coincide on $\mathfrak{a} \cap \mathfrak{h}$. Finally since ν is an automorphism of $\mathfrak{g}$, it is obvious that λ' is a root of $\mathfrak{g}$ with respect to $\mathfrak{a}$ if and only if λ is a root with respect to $\mathfrak{h}$. Hence if we identify linear functions on $\mathfrak{h}$ with those on $\mathfrak{a}$ under the mapping $\lambda \to \lambda'$, the two sets of roots coincide. Then Σ is exactly the set of those roots α for which $H_\alpha \notin \mathfrak{a}_{\mathfrak{k}}$ (in the notation of Section 6). Let Q be the set of those roots in P which are not identically zero on $\nu(\mathfrak{a}_{\mathfrak{p}})$. Then it follows from Lemma 16 that Q is the disjoint union of C_i, C_{ij}, P_0, P_i, P_{ij} $(1 \leqq i < j \leqq s)$. Moreover it is obvious that

$$|D(\exp H)|^{\frac{1}{2}} = \Big|\prod_{\alpha \in Q} (e^{\alpha(H)} - e^{-\alpha(H)})\Big| \qquad (H \in \mathfrak{a}_{\mathfrak{p}}).$$

Now put

$$H = \sum_{i=1}^{s} t_i (X_{\gamma_i} + X_{-\gamma_i}) \qquad (t_i \in C).$$

Then

$$\nu(H) = \sum_{i=1}^{s} t_i H_{\gamma_i} \text{ and therefore}$$

$$\alpha(H) = \alpha(\nu(H)) = \sum_{i=1}^{s} t_i \alpha(H_{\gamma_i}) \qquad (\alpha \in Q).$$

Since $\gamma_i(H_{\gamma_j}) = 2\delta_{ij}$ $(1 \leqq i, j \leqq s)$ it is obvious (see Section 6) that

$$\begin{aligned}
&\alpha(H) = -t_i \quad \text{if} \quad \alpha \in C_i,\\
&\alpha(H) = t_j - t_i \quad \text{if} \quad \alpha \in C_{ij},\\
&\alpha(H) = t_i \quad \text{if} \quad \alpha \in P_i,\\
&\alpha(H) = t_j + t_i \quad \text{if} \quad \alpha \in P_{ij} \qquad (1 \leqq i < j \leqq s).
\end{aligned}$$

Put

$$Q_1 = \bigcup_{1 \leqq i \leqq s} (C_i \cup P_i), \qquad Q_2 = \bigcup_{1 \leqq i < j \leqq s} (C_{ij} \cup P_{ij}),$$

Then in the notation of Section 6,

$$\begin{aligned}
&\Big| \prod_{\alpha \in Q_1} (e^{\alpha(H)} - e^{-\alpha(H)}) \Big| = \prod_{1 \leqq i \leqq s} 2^{2r_i} |\sinh t_i|^{2r_i},\\
&\Big| \prod_{\alpha \in P_0} (e^{\alpha(H)} - e^{-\alpha(H)}) \Big| = \prod_{1 \leqq i \leqq s} |4 \sinh t_i \cosh t_i|.
\end{aligned}$$

Moreover since

$$\sinh(t_j - t_i) \sinh(t_j + t_i) = (\cosh t_j)^2 - (\cosh t_i)^2,$$

it follows that

$$\Big| \prod_{\alpha \in Q_2} (e^{\alpha(H)} - e^{-\alpha(H)}) \Big| = \prod_{1 \leqq i < j \leqq s} 2^{2r_{ij}} |(\cosh t_j)^2 - (\cosh t_i)^2|^{r_{ij}}.$$

Hence

$$\begin{aligned}
|D(\exp H)|^{\frac{1}{2}} = &\prod_{1 \leqq i \leqq s} 2^{2(r_i+1)} |\sinh t_i|^{2r_i+1} |\cosh t_i|\\
&\times \prod_{1 \leqq i < j \leqq s} 2^{2r_{ij}} |(\cosh t_j)^2 - (\cosh t_i)^2|^{r_{ij}}.
\end{aligned}$$

Now suppose $\cosh t_i \neq 0$ $1 \leqq i \leqq s$ and put[6] $t_i' = \log(\cosh t_i)$. Let

$$H' = \sum_{i=1}^{s} t_i H_{\gamma_i}$$

and consider the expression

$$\Delta(2H') = \prod_{\alpha \in C'} (e^{\alpha(H')} - e^{-\alpha(H')})$$

where C' is the set of all positive compact roots which do not vanish identically on $\nu(\mathfrak{a}_0)$. It follows from Lemma 16 that C' is the disjoint union of C_i and C_{ij} $(1 \leqq i < j \leqq s)$. Therefore it is clear that

$$\Delta(2H') = \prod_i \{\cosh t_i - (1/\cosh t_i)\} \prod_{i<j} \{(\cosh t_j/\cosh t_i) - (\cosh t_i/\cosh t_j)\}^{r_{ij}}$$

$$= \prod_i \{(\sinh t_i)^{2r_i}/(\cosh t_i)^{r_i}\} \prod_{i<j} \{(\cosh t_j)^2 - (\cosh t_i)^2\}^{r_{ij}}/(\cosh t_i \cosh t_j)^{r_{ij}}.$$

Hence

$$|D(\exp H)|^{\frac{1}{2}} = |\Delta(2H')| \{\prod_i 2^{2r_i+2} |\cosh t_i|^{r_i+1} |\sinh t_i|\}$$

$$\times \prod_{i<j} |4 \cosh t_i \cosh t_j|^{r_{ij}}.$$

But we know from Lemma 18 that

$$r_i + \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji} = 2\rho_+(H_i) - 2$$

Therefore

$$\{\prod_i 2^{2r_i+2} |\cosh t_i|^{r_i+1}\} \prod_{i<j} |4 \cosh t_i \cosh t_j|^{r_{ij}}$$

$$= \prod_i 2^{2\rho_+(H\gamma_i)} |\cosh t_i|^{2\rho_+(H\gamma_i)-1} = 2^p \prod_i |\cosh t_i|^{2\rho_i(H\gamma_i)-1}$$

where $p = \sum_{1\leqq i\leqq s} 2\rho_+(H_{\gamma_i})$. Thus we have the following result.

$$|D(\exp H)|^{\frac{1}{2}} = 2^p |\Delta(2H')| \prod_i \{|\cosh t_i|^{2\rho_+(H\gamma_i)-1} |\sinh t_i|\}$$

We now introduce the partial order in the set $P_0 = (\gamma_1, \cdots, \gamma_s)$ as described at the end of Section 6. Let $\beta_1, \cdots, \beta_r$ be all the (distinct) minimal elements in P_0 under this order. For any i $(1 \leqq i \leqq r)$ consider the set σ_i of all $\gamma \in P_0$ such that $\frac{1}{2}(\gamma - \beta_i)$ is a root. Then γ, γ' are two distinct elements of σ_i, it follows from Lemma 17 that either $\gamma \prec \gamma'$ or $\gamma' \prec \gamma$. This shows that σ_i is simply ordered. Therefore we may write it in the form

$$\beta_i = \beta_{i1} \prec \beta_{i2} \cdots \prec \beta_{is_i}.$$

Moreover if $i \neq j$, σ_i, σ_j are disjoint $(1 \leqq i, j \leqq r)$. For otherwise suppose $\gamma \in \sigma_i \cap \sigma_j$. Then $\frac{1}{2}(\gamma - \beta_i)$ and $\frac{1}{2}(\gamma - \beta_j)$ are both roots and therefore again from Lemma 17, $\frac{1}{2}(\beta_i - \beta_j)$ is a root. This however is impossible since both β_i and β_j are minimal in P_0. Thus P_0 is the disjoint union of $\sigma_1, \cdots, \sigma_r$. We put $t_{ij} = t_k$ if $\beta_{ij} = \gamma_k$ $(1 \leqq i \leqq r, 1 \leqq j \leqq s_i, 1 \leqq k \leqq s)$. Let $\mathfrak{b}$ denote the subset of $\mathfrak{a}_{\mathfrak{p}_0}$ consisting of all elements $H = t_1 H_{\gamma_1} + \cdots + t_s H_{\gamma_s}$ $(t_j \in R)$ such that

$$0 < t_{i1} < t_{i2} < \cdots < t_{is_i} \qquad (1 \leqq i \leqq r).$$

Similarly let $\mathfrak{b}^*$ denote the subset of $\mathfrak{b}$ consisting of those H which satisfy the additional condition $t_{is_i} < \frac{\pi}{2}$ $(1 \leqq i \leqq r)$. Then we have the following result.

LEMMA 23. *Put* $e(X) = \exp(-1)^{\frac{1}{2}}X$ $(X \in \mathfrak{g})$ *and*

$$\Delta(H) = \prod_{\alpha \in C'} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)}) \qquad (H \in \mathfrak{h}).$$

For any $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i})$ *in* $\mathfrak{b}$ *let* H' *denote the element*[6]

$$H' = \log(\cosh t_1)H_{\gamma_1} + \cdots + \log(\cosh t_s)H_{\gamma_s}$$

in $\mathfrak{h}$. *Similarly for any* $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i})$ *in* $\mathfrak{b}^*$, *let* H^* *denote the element*

$$H^* = \log(\cos t_1)H_{\gamma_1} + \cdots + \log(\cos t_s)H_{\gamma_s}.$$

Then

$$|D(\exp H)|^{\frac{1}{2}} = 2^p \Delta(2H') \prod_i (\cosh t_i)^{2\rho_+(H\gamma_i)-1} \prod_i \sinh t_i \qquad (H \in \mathfrak{b})$$

and

$$|D(e(H))|^{\frac{1}{2}} = (-1)^p 2^p \Delta(2H^*) \prod_i (\cos t_i)^{2\rho_+(H\gamma_i)-1} \prod_i \sin t_i \qquad (H \in \mathfrak{b}^*).$$

If $H \in \mathfrak{b}$, $t_i > 0$ $(1 \leqq i \leqq s)$ and therefore $\sinh t_i > 0$. Hence in view of our earlier result, in this case it is enough to prove that $\Delta(2H')$ is real and positive. But we have already seen that

$$\Delta(2H') = \prod_i \{(\sinh t_i)^{2r_i}/(\cosh t_i)^{r_i}\} \times \prod_{i<j} \{(\cosh t_j)^2 - (\cosh t_i)^2\}^{r_{ij}}/(\cosh t_i \cosh t_j)^{r_{ij}}.$$

Since $\cosh t$ is a positive increasing function of t for $t > 0$, $\Delta(2H')$ has the same sign as

$$\eta = \prod_{1 \leqq i < j \leqq s} (t_j - t_i)^{r_{ij}}.$$

Now if $\frac{1}{2}(\gamma_j - \gamma_i)$ is not a root, we know from Lemma 19 that r_{ij} is even. On the other hand if $\frac{1}{2}(\gamma_j - \gamma_i)$ is a root, $\gamma_i \prec \gamma_j$ and therefore, in view of the definition of $\mathfrak{b}$, $t_j - t_i > 0$. This shows that $\Delta(2H') \geqq 0$ and so the first assertion of the lemma follows.

Now we come to the second case when $H \in \mathfrak{b}^*$. Since $\cosh((-1)^{\frac{1}{2}}t) = \cos t$ and $\sinh((-1)^{\frac{1}{2}}t) = (-1)^{\frac{1}{2}}\sin t$ $(t \in R)$, it follows that

$$\Delta(2H^*) = \prod_i (-1)^{r_i}\{(\sin t_i)^{2r_i}/(\cos t_i)^{r_i}\} \times \prod_{i<j} \{(\cos t_j)^2 - (\cos t_i)^2\}^{r_{ij}}/(\cos t_i \cos t_j)^{r_{ij}}$$

and therefore $\Delta(2H^*)$ is real. Moreover since $\cos t$ is a positive decreasing function of t in the interval $0 < t < \frac{\pi}{2}$ it follows that $\Delta(2H^*)$ has the same sign as

$$\prod_i (-1)^{r_i} \prod_{i<j} (t_i - t_j)^{r_{ij}} = (-1)^q \eta$$

where $q = \sum_i r_i + \sum_{i<j} r_{ij}$. We have seen above that $\eta \geqq 0$ on $\mathfrak{b}$ and therefore also on $\mathfrak{b}^*$ and from Lemma 18,

$$q = \sum_{i=1}^{s} \{2\rho_+(H_{\gamma_i}) - 2\} = p - 2s.$$

Hence $(-1)^q = (-1)^p$. On the other hand $\cos t$ and $\sin t$ are both positive on the interval $0 < t < \frac{1}{2}$ and so the second statement of the lemma is an immediate consequence of our earlier expression for $|D|^{\frac{1}{2}}$.

Let B and B^* be the images in $\mathfrak{A}$ and $\mathfrak{A}^*$ of $\mathfrak{b}$ and $\mathfrak{b}^*$ under the mappings $H \to \exp H$ and $H \to e(H)$ respectively.

LEMMA 24. *$B \cap \mathfrak{A}'$ is both open and closed in $\mathfrak{A}'$. Similarly $B^* \cap \mathfrak{A}^{*\prime}$ is both open and closed in $\mathfrak{A}^{*\prime}$.*

Since $\mathfrak{b}$ and $\mathfrak{b}^*$ are obviously open in $\mathfrak{a}_{\mathfrak{p}_0}$ and since the mappings $H \to \exp H$ and $H \to e(H)$ are regular on $\mathfrak{a}_{\mathfrak{p}_0}$, it follows that B and B^* are open in $\mathfrak{A}$ and $\mathfrak{A}^*$ respectively. Let $\bar{\mathfrak{b}}$ and $\bar{\mathfrak{b}}^*$ respectively denote the closures of $\mathfrak{b}$ and $\mathfrak{b}^*$ in the real Euclidean space $\mathfrak{a}_{\mathfrak{p}_0}$. Then $\mathfrak{b}^*$ is compact. Since $H \to \exp H$ is a topological mapping of $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$ (see Section 12), the image $\exp \bar{\mathfrak{b}}$ of $\bar{\mathfrak{b}}$ under this mapping is closed in $\mathfrak{A}$. Also $e(\bar{\mathfrak{b}}^*)$ is compact and therefore closed in $\mathfrak{A}^*$. Hence it is enough to prove that

$$(\exp \bar{\mathfrak{b}}) \cap \mathfrak{A}' = B \cap \mathfrak{A}', \qquad e(\bar{\mathfrak{b}}^*) \cap \mathfrak{A}^{*\prime} = B^* \cap \mathfrak{A}^{*\prime}.$$

Now let $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \in R)$ be a point in $\bar{\mathfrak{b}}$. Then $0 \leqq t_i$ and if $\gamma_i \prec \gamma_j$, $t_i \leqq t_j$. On the other hand we have seen that

$$|D(\exp H)|^{\frac{1}{2}} = \prod_i 2^{2(r_i+1)} |\sinh t_i|^{2r_i+1} |\cosh t_i| \times \prod_{i<j} 2^{2r_{ij}} |(\cosh t_j)^2 - (\cosh t_i)^2|^{r_{ij}}.$$

Hence if $D(\exp H) \neq 0$, $t_i \neq 0$ and $t_i \neq t_j$ if $r_{ij} > 0$ $(1 \leqq i < j \leqq s)$. This proves that in this case $H \in \mathfrak{b}$ and therefore $(\exp \bar{\mathfrak{b}}) \cap \mathfrak{A}' = B \cap \mathfrak{A}'$. Now suppose H lies in $\bar{\mathfrak{b}}^*$. Then we have the additional conditions $0 \leqq t_i \leqq \frac{\pi}{2}$ $(i = 1, \cdots, s)$. Moreover again we know that

$$|D(e(H))|^{\frac{1}{2}} = \prod_i 2^{2(r_i+1)} |\sin t_i|^{2r_i+1} |\cos t_i| \times \prod_{i<j} 2^{2r_{ij}} |(\cos t_j)^2 - (\cos t_i)^2|^{r_{ij}}$$

and therefore if $D(e(H)) \neq 0$, $t_i \neq 0$, $\frac{\pi}{2}$ and $t_i \neq t_j$ whenever $r_{ij} > 0$

$(1 \leqq i < j \leqq s)$. This shows that $H \varepsilon \mathfrak{b}^*$ in this case and therefore

$$e(\mathfrak{b}) \cap \mathfrak{A}^{*\prime} = B^* \cap \mathfrak{A}^{*\prime}.$$

COROLLARY. *Every connected component of $B \cap \mathfrak{A}'$ or $B^* \cap \mathfrak{A}^{*\prime}$ is also a connected component of $\mathfrak{A}'$ or $\mathfrak{A}^{*\prime}$ respectively. The number of connected components of $B \cap \mathfrak{A}'$ is the same as that of $B^* \cap \mathfrak{A}^{*\prime}$.*

The first statement is obvious from the above lemma. Now first we claim that the mapping $H \to e(H)$ is univalent on $\mathfrak{b}^*$. For suppose $e(H_1) = e(H_2)$ $(H_1, H_2 \varepsilon \mathfrak{b}^*)$. Then it is obvious that if α is any root in Σ, $\alpha(H_1) - \alpha(H_2)$ must be an integral multiple of 2π. Hence in particular $\gamma_j(v(H_1)) - \gamma_j(v(H_2)) = 2n_j\pi$ $(1 \leqq j \leqq s)$ where n_j is an integer. But since $H_1, H_2 \varepsilon \mathfrak{b}^*$, $0 < \gamma_j(v(H_i)) < \pi$ $i = 1, 2$ and therefore $n_j = 0$ $(1 \leqq j \leqq s)$. This however implies that $H_1 = H_2$. Since we already know that the mapping $H \to e(H)$ of $\mathfrak{b}^*$ onto B^* is open and continuous it follows that it is topological.

Now let σ_α $(\alpha \varepsilon \Sigma)$ denote the hyperplane in $\mathfrak{a}_{\mathfrak{p}_0}$ consisting of all points H such that $\alpha(H) = 0$. Let $\mathfrak{a}'_{\mathfrak{p}_0}$ denote the complement of $\bigcup_{\alpha \varepsilon \Sigma} \sigma_\alpha$ in $\mathfrak{a}_{\mathfrak{p}_0}$. Then it is obvious that $D(\exp H) \neq 0$ $(H \varepsilon \mathfrak{a}_{\mathfrak{p}_0})$ if and only if $H \varepsilon \mathfrak{a}'_{\mathfrak{p}_0}$. Since the exponential mapping of $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$ is topological, $B \cap \mathfrak{A}'$ and $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$ have the same number of components. Let $\mathfrak{b}_0$ be a connected component of $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Then it is obvious that every root $\alpha \varepsilon \Sigma$ must keep constant sign on $\mathfrak{b}_0$. Therefore since $\mathfrak{b}$ is obviously a convex set, the same holds for $\mathfrak{b}_0$. Moreover if $H \varepsilon \mathfrak{b}_0$, it is obvious that the half-line consisting of all points tH $(t > 0)$ lies entirely in $\mathfrak{b}_0$. But if t is sufficiently small and positive $tH \varepsilon \mathfrak{b}^*$. This implies that $\mathfrak{b}_0$ contains some connected component of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Also if H_1, H_2 are two points in $\mathfrak{b}^*$ which both lie in $\mathfrak{b}_0$, the straight line-segment J joining them is also contained in $\mathfrak{b}_0$. Hence $J \subset \mathfrak{a}'_{\mathfrak{p}_0}$. But since $\mathfrak{b}^*$ is obviously convex, $J \subset \mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$. This shows that $\mathfrak{b}_0$ cannot contain two distinct connected components of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ and therefore every component of $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$ contains exactly one component of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Conversely it is obvious that every component of $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ is contained in exactly one component of $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$. Hence these two sets have the same number of components.

On the other hand if $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ lies in $\mathfrak{b}^*$,

$$|D(e(H))|^{\frac{1}{2}} = \prod_i 2^{2(r_i+1)} |\sin t_i|^{2r_i+1} |\cos t_i| \times \prod_{i<j} 2^{2r_{ij}} |(\cos t_j)^2 - (\cos t_i)^2|^{r_{ij}}.$$

Since $0 < t_i < \frac{\pi}{2}$, $D(e(H)) = 0$ if and only if $t_i = t_j$ for some pair of indices i, j $(i < j)$ with $r_{ij} > 0$. But clearly this happens if and only if $\alpha(H) = 0$ for some $\alpha \in \Sigma$. Hence $D(e(H)) \neq 0$ if and only if $H \in \mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ and so the mapping $H \to e(H)$ maps $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$ topologically onto $B^* \cap \mathfrak{A}^{*\prime}$. Similarly $H \to \exp H$ $(H \in \mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0})$ maps $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$ topologically onto $B \cap \mathfrak{A}'$. Therefore $B \cap \mathfrak{A}'$, $\mathfrak{b} \cap \mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{b}^* \cap \mathfrak{a}'_{\mathfrak{p}_0}$, $B^* \cap \mathfrak{A}^{*\prime}$ all have the same number of components.

9. Application to representations of G. Let G be the simply connected covering group of G_0 and let K be the analytic subgroup of G corresponding to $\mathfrak{k}_0$. Define the complex manifold W containing G as in [5(f), § 3] and the mapping Γ of G into the center $\mathfrak{c}$ of $\mathfrak{k}$ as in [5(f), § 6]. Then $\Gamma(xu) = \Gamma(ux) = \Gamma(u) + \Gamma(x)$ $(u \in K, x \in G)$ [5(f), Lemma 13] and if λ is a real linear function on $\mathfrak{h}$ (see [5(e), § 2]) $\lambda(\Gamma(u))$ is pure imaginary [5(f), Lemma 23]. Let Λ be a real linear function on $\mathfrak{h}$ such that $\Lambda(H_\alpha)$ is a non-negative integer for every positive compact root α. Consider a fundamental system $(\alpha_1, \cdots, \alpha_l)$ of positive roots and suppose that $(\alpha_1, \cdots, \alpha_m)$ are all the totally positive roots in this system. Let Λ_0 denote the linear function on $\mathfrak{h}$ given by $\Lambda_0(H_{\alpha_i}) = 0$ $1 \leqq i \leqq m$ and $\Lambda_0(H_{\alpha_i}) = \Lambda(H_{\alpha_i})$ $m < i \leqq l$. Then we can define (see [5(f), § 6]) an irreducible complex representation σ of G_c on a finite-dimensional Hilbert space V such that σ is unitary on U and its highest weight is Λ_0. Let ϕ_0 be a unit vector in V belonging to the highest weight Λ_0 and put $\lambda = \Lambda - \Lambda_0$. Then if $x \to \bar{x}$ $(x \in G)$ denotes the natural mapping of G onto G_0, we consider the function

$$\psi_\Lambda(x) = (\phi_0, \sigma(\bar{x})\phi_0)\, e^{\lambda(\Gamma(x))} \qquad (x \in G)$$

on G. Since the center of G lies in K and Λ is real, $|\psi_\Lambda(x)|$ depends only on $\bar{x}$. We propose to consider the integral (cf. [5(f), § 9])

$$\int_{G_0} |\psi_\Lambda(x)|^2\, d\bar{x}$$

where $d\bar{x}$ is the Haar measure on G_0. For any $a \in \mathfrak{A}$ let $\log a$ denote the unique element $H \in \mathfrak{a}_{\mathfrak{p}_0}$ such that $a = \exp H$. We put $\Gamma(a) = \Gamma(\exp' \log a)$ where $X \to \exp' X$ $(X \in \mathfrak{g}_0)$ is the exponential mapping of $\mathfrak{g}_0$ into G. Then if we normalize the various Haar measures in accordance with Lemma 22 and take into account the fact that $\lambda(\Gamma(a))$ is real for $a \in \mathfrak{A}$ [5(f), Lemma 23], it follows that

$$\int_{G_0} |\psi_\Lambda(x)|^2\, d\bar{x} = wN^{-1} \int_B e^{2\lambda(\Gamma(a))}\, |D(a)|^{\frac{1}{2}}\, da \int_{K_0 \times K_0} (\phi_0, \sigma(kak')\phi_0)^2\, dk dk'$$

where B is defined[8] as in Lemma 24 and N is the number of connected components of $B \cap \mathfrak{A}'$. Thus we are led to the integral

$$\int_{K_0 \times K_0} |(\phi^*, \sigma(kak')\phi_0)|^2 \, dkdk' \qquad (a \in B).$$

Let $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i}) \in \mathfrak{b}$ $(t_i \in R)$ and let $a = \exp H$. Then we know from Lemma 20 that

$$a = \zeta h(a) \xi$$

where $\zeta \in \mathfrak{P}_c^-$, $\xi \in \mathfrak{P}_c^+$ (in the notation of Section 7) and

$$h(a) = \exp\Big(\sum_{i=1}^{\delta} \log(\cosh t_i) H_i\Big).$$

If we denote the corresponding representation of $\mathfrak{g}$ also by σ, it is obvious that $\sigma(\mathfrak{p}_+)\phi_0 = 0$ and therefore $\sigma(p)\phi_0 = \phi_0$ for $p \in \mathfrak{P}_c^+$. Since $k\mathfrak{P}_c^- k^{-1} = \mathfrak{P}_c^-$, $k\mathfrak{P}_c^+ k^{-1} = \mathfrak{P}_c^+ (k \in K_c)$ and $\tilde{\theta}(\mathfrak{P}_c^-) = \mathfrak{P}_c^+$, it follows that

$$(\phi_0, \sigma(kak')\phi_0) = (\phi_0, \sigma(kh(a)k')\phi_0).$$

Let A_c be the complex analytic subgroup of G_c corresponding to $\mathfrak{h}$. Then $h(a) \in A_c \subset K_c$ and

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)|^2 \, dkdk' = \int_{K_0 \times K_0} |(\phi_0, \sigma(kh(a)k')\phi_0)|^2 \, dkdk'.$$

On the other hand $\sigma(H)\phi_0 = \Lambda_0(H)\phi_0$ $(H \in \mathfrak{h})$ and $\sigma(X_\alpha)\phi_0 = 0$ for any positive root α. Since this holds in particular for every positive compact root, it follows from Lemma 2 of [5(e)] (applied to $\mathfrak{k}$) that the subspace V_0 of V spanned by $\sigma(k)\phi_0$ $(k \in K_c)$ is irreducible under K_c. Let σ_0 denote the corresponding representation of K_c (and k) on V_0. Then obviously Λ_0 is the highest weight of σ_0. Now if we make use of the Schur orthogonality relations for the irreducible representation of the compact group K_0 on V_0, we find easily that

$$\begin{aligned}\int_{K_0 \times K_0} |(\phi_0, \sigma(kyk')\phi_0)|^2 \, dkdk' &= (\dim V_0)^{-1} \int_{K_0} |\sigma(yk)\phi_0|^2 \, dk \\ &= (\dim V_0)^{-2} \operatorname{Sp} \sigma_0(\tilde{\theta}(y)y)\end{aligned}$$

if $y \in K_c$. Hence

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)|^2 \, dkdk' = (\dim V_0)^{-2} \chi_{\Lambda_0}(h(a))^2)$$

[8] Here we have to make use of the obvious fact that the complement of $\mathfrak{A}'$ in $\mathfrak{A}$ is of measure zero with respect to da.

where χ_{Λ_0} is the character of σ_0. But if $P_\mathfrak{k}$ is the set of all compact positive roots and $\rho_\mathfrak{k} = \frac{1}{2}\sum_{\alpha \in P_\mathfrak{k}} \alpha$, we know (see Weyl [11(a)]) that

$$\dim V_0 = \prod_{\alpha \in P_\mathfrak{k}} \{\Lambda_0(H_\alpha) + \rho_\mathfrak{k}(H_\alpha)/\rho_\mathfrak{k}(H_\alpha)\}$$
$$= \prod_{\alpha \in P_\mathfrak{k}} \{\Lambda(H_\alpha) + \rho_\mathfrak{k}(H_\alpha)/\rho_\mathfrak{k}(H_\alpha)\}$$

since $\lambda(H_\alpha) = 0$ $(\alpha \in P_\mathfrak{k})$ from Lemma 13 of [5(e)].

We now claim that

$$\lambda(\Gamma(a)) = \sum_{i=1}^{s} \log(\cosh t_i)\lambda(H_{\gamma_i})$$

where (in accordance with our convention) $\log(\cosh t_i)$ is real $(1 \leqq i \leqq s)$. Since both sides are linear in λ. it is enough to prove this under the assumption that $\lambda(H_{\alpha_i})$ $1 \leqq i \leqq m$ are all non-negative integers and $\lambda(H_{\alpha_i}) = 0$ $m < i \leqq l$ (see the proof of Lemma 23 of [5(f)]). But then we can find an irreducible complex representation σ' of G_c on a finite-dimensional Hilbert space V' with the highest weight λ and assume that σ' is unitary on U (see [5(f), § 6]). Let ϕ' be a unit vector in V' belonging to the weight λ. Then, as we have seen during the proof of Lemma 23 of [5(f)],

$$(\phi', \sigma'(\tilde{x})\phi') = e^{\lambda(\Gamma(a))} \qquad (x \in G)$$

and therefore by the argument which we have already used above,

$$e^{\lambda(\Gamma(a))} = (\phi', \sigma'(a)\phi') = (\phi', \sigma'(h(a))\phi')$$
$$= \exp\left(\sum_{i=1}^{s} (\log\cosh t_i)\lambda(H_{\gamma_i})\right)$$

since ϕ' belongs to the weight λ. Our assertion now follows from the fact that both $\lambda(\Gamma(a))$ and $\sum_i (\log\cosh t_i)\lambda(H_{\gamma_i})$ are real.

If b and μ are two real numbers and b is positive we define b^μ in the usual way by $b^\mu = \exp(\mu \log b)$. Then the above result shows that

$$e^{\lambda(\Gamma(a))} = \prod_{i=1}^{s} (\cosh t_i)^{2\lambda(H_{\gamma_i})}$$

For any point $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \in R)$ in $\mathfrak{a}_{\mathfrak{p}_0}$, we regard $(t_1, \cdots, t_s)$ as the coordinates of H and denote by dt the measure $dt_1 dt_2 \cdots dt_s$ on $\mathfrak{a}_{\mathfrak{p}_0}$. It is obvious that $dH = c\,dt$ where c is a positive constant and dH is the element of volume corresponding to the Euclidean metric on $\mathfrak{a}_{\mathfrak{p}_0}$ (see Section 8). Hence it follows from the above remarks that

$$\int_{G_0} |\psi_\Lambda(x)|^2\, d\tilde{x} = wN^{-1}(\dim V_0)^{-1} \int_B \chi_{\Lambda_0}((h(a))^2) e^{2\lambda(\Gamma(a))} |D(a)|^{\frac{1}{2}}\, da$$

$$= wcN^{-1}(\dim V_0)^{-2} \int_{\mathfrak{b}} \chi_{\Lambda_0}(\exp 2H') \prod_{i=1}^{s} (\cosh t_i)^{2\lambda(H\gamma_i)} |D(\exp H)|^{\frac{1}{2}}\, dt$$

where $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ and

$$H' = \sum_{i=1}^{s} (\log \cosh t_i) H_{\gamma_i} \qquad (t_i \in R).$$

Let $\mathfrak{w}_\mathfrak{k}$ be the subgroup of the Weyl group (of $\mathfrak{g}$ with respect to $\mathfrak{h}$) generated by the Weyl reflexions s_α corresponding to $\alpha \in P_\mathfrak{k}$. For any $s \in \mathfrak{w}_\mathfrak{k}$ we define $\epsilon(s) = 1$ or -1 according as the permutation $\alpha \to s\alpha$ of the set of all compact roots α, is even or odd. As in Lemma 16, let C_0 denote the set of those positive compact roots which vanish identically on $\nu(\mathfrak{a}_\mathfrak{p})$ and let C' be the complement of C_0 in $P_\mathfrak{k}$. Then if $\rho_\mathfrak{k} = \frac{1}{2}\sum_{\alpha \in P_\mathfrak{k}} \alpha$ and $\rho_0 = \frac{1}{2}\sum_{\alpha \in C_0} \alpha$, we have the following result.

LEMMA 25. *Put* $\Lambda''_0 = \Lambda_0 + \rho_\mathfrak{k}$. *Then if* w_0 *is the order of the subgroup* $\mathfrak{w}_\mathfrak{k}^0$ *of* $\mathfrak{w}_\mathfrak{k}$ *generated by* s_α $(\alpha \in C_0)$,

$$\chi_{\Lambda_0}(\exp H) \prod_{\alpha \in C'} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})$$

$$= \{w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha)\}^{-1} \sum_{s \in \mathfrak{w}_\mathfrak{k}} \epsilon(s) \{\prod_{\alpha \in C_0} s\Lambda''_0(H_\alpha)\} e^{s\Lambda''_0(H)}$$

for all $H \in \nu(\mathfrak{a}_\mathfrak{p})$.

For each $\alpha \in C'$ we define a holomorphic linear differential operator D_α of order one on the complex Euclidean space $\mathfrak{h}$ such that $D_\alpha \mu = \mu(H_\alpha)$ for every linear function μ on $\mathfrak{h}$. Then the operators D_α obviously commute with each other. Put $D = \prod_{\alpha \in C'} D_\alpha$. We know (see Weyl [11(a)]) that

$$\chi_{\Lambda_0}(\exp H \prod_{\alpha \in C'} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)}) \prod_{\alpha \in C_0} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)})$$

$$= \sum_{s \in \mathfrak{w}_\mathfrak{k}} \epsilon(s) e^{s\Lambda''_0(H)} \qquad (H \in \mathfrak{h}).$$

Hence applying the differential operator D to both sides and evaluating the result at a point H in $\nu(\mathfrak{a}_\mathfrak{p})$, we find that

$$\chi_{\Lambda_0}(\exp H) \prod_{\alpha \in C'} (e^{\frac{1}{2}\alpha(H)} - e^{-\frac{1}{2}\alpha(H)}) (D\Delta_0)_H$$

$$= \sum_{s \in \mathfrak{w}_\mathfrak{k}} \epsilon(s) \{\prod_{\alpha \in C_0} s\Lambda''_0(H_\alpha)\} e^{s\Lambda''_0(H)} \qquad H \in \nu(\mathfrak{a}_\mathfrak{p}))$$

since $\alpha(H) = 0$ of $\alpha \in C_0$ and $H \in \nu(\mathfrak{a}_\mathfrak{p})$. Here

$$\Delta_0 = \prod_{\alpha \in C_0} (e^{\alpha/2} - e^{-\alpha/2})$$

and $(D\Delta_0)_H$ denotes the value of $D\Delta_0$ at H. But it follows from well-known arguments (Weyl [11(a)]) that[9]

$$\Delta_0 = \sum_{s \in \mathfrak{w}_\mathfrak{k}^0} \epsilon_0(s) e^{s\rho_0}$$

where $\epsilon_0(s) = \pm 1$ and is determined by the rule

$$\prod_{\alpha \in C_0} (e^{\frac{1}{2}s\alpha} - e^{-\frac{1}{2}s\alpha}) = \epsilon_0(s)\Delta_0 \qquad (s \in \mathfrak{w}_\mathfrak{k}^0).$$

Hence

$$D\Delta_0 = \sum_{s \in \mathfrak{w}_\mathfrak{k}^0} \epsilon_0(s) \{ \prod_{\alpha \in C_0} s\rho_0(H_\alpha) \} e^{s\rho_0}.$$

But it is obvious that $\epsilon_0(s) = (-1)^q$ $(s \in \mathfrak{w}_\mathfrak{k}^0)$ if q is the number of negative roots among $s\alpha$ $(\alpha \in C_0)$. Therefore

$$\prod_{\alpha \in C_0} s\rho_0(H_\alpha) = \epsilon_0(s) \prod_{\alpha \in C_0} \rho_0(H_\alpha)$$

and

$$(D\Delta_0)_H = w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha) \qquad (H \in \nu(\mathfrak{a}_\mathfrak{p}))$$

since $\rho_0(H) = 0$. Moreover if $\beta \in C_0$, not all the roots $s_\beta\alpha (\alpha \in C_0)$ are positive since $s_\beta\beta = -\beta$. This shows that $\sum_{\alpha \in C_0} s_\beta\alpha < \sum_{\alpha \in C_0} \alpha$ and therefore $s_\beta\rho_0 < \rho_0$. Since this implies that $\rho_0(H_\beta) > 0$, the lemma now follows.

Now again for any point $H = \sum_{i=1}^{s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \in R)$ in $\mathfrak{b}$ put $H' = \sum_i \log(\cosh t_i) H_{\gamma_i}$. Then if $\rho = \frac{1}{2} \sum_{\alpha \in P} \alpha$ and $\Lambda' = \Lambda + \rho$, $\Lambda' = \Lambda''_0 + \lambda + \rho_+$. Moreover since $\lambda(H_\alpha) = 0$ for every compact root α [5(e), Lemma 13], it follows that $s\lambda = \lambda$ for all $s \in \mathfrak{w}_\mathfrak{k}$. Similarly it follows from Lemma 10 of [5(e)] that $s\rho_+ = \rho_+$ $(s \in \mathfrak{w}_\mathfrak{k})$ and therefore $\rho_+(H_\alpha) = 0$ if $\alpha \in P_\mathfrak{k}$. Hence we conclude from Lemmas 23 and 25 that

$$\chi_{\Lambda_0}(\exp 2H') \prod_{i=1}^{s} (\cosh t_i)^{2\lambda(H\gamma_i)} \mid D(\exp H) \mid^{\frac{1}{2}}$$
$$= 2^p \{ w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha) \}^{-1} \prod_i (\sinh t_i)$$
$$\times \sum_{\tau \in \mathfrak{w}_\mathfrak{k}} \epsilon(\tau) \{ \prod_{\alpha \in C_0} \tau\Lambda'(H_\alpha) \} \prod_i (\cosh t_i)^{2\tau\Lambda'(H\gamma_i)-1}.$$

---

[9] We have to apply here Weyl's result to the subalgebra $\mathfrak{a}_\mathfrak{k} + \sum_{\alpha \in C_0} (CX_\alpha + CX_{-\alpha})$ of $\mathfrak{g}$ which is obviously reductive [7].

Put $y_i = (\cosh t_i)^{-1}$. Then $\mathfrak{b}$ corresponds to the region defined by $0 < y_i < 1$ and $y_i > y_j$ if $\gamma_i \prec \gamma_j$ $(1 \leqq i < j \leqq s)$. Let us denote this region by $\mathfrak{b}_y$. Then if $dy = dy_1 dy_2 \cdots dy_s$, it is clear that

$$\int_{\mathfrak{b}} \chi_{\lambda_0}(\exp 2H')\, e^{2\lambda(\Gamma(\exp H))} \mid D(\exp H) \mid^{\frac{1}{2}} dt$$
$$= 2^p \{ w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha) \}^{-1} \int_{\mathfrak{b}_y} \{ \sum_{\tau \in \mathfrak{W}_l} \epsilon(\tau) \Big(\prod_{\alpha \in C_0} \tau\Lambda'(H_\alpha) \Big) \prod_i y_i^{-2\tau\Lambda'(H\gamma_i)-1} \} dy.$$

Put $y_{ij} = (\cosh t_{ij})^{-1}$ $(1 \leqq i \leqq r, 1 \leqq j \leqq s_i)$ in the notation of Section 8. Then the region $\mathfrak{b}_y$ is defined by the inequalities

$$1 > y_{i1} > y_{i2} > \cdots > y_{is_i} > 0 \qquad (1 \leqq i \leqq r)$$

and an elementary computation shows that if q_{ij} are *positive* real numbers

$$\int_{\mathfrak{b}_y} \prod_{i,j} y_{ij}^{q_{ij}-1} dy = \prod_{1 \leqq i \leqq r} \prod_{1 \leqq k \leqq s_i} \Big(\sum_{k \leqq j \leqq s_i} q_{ij} \Big)^{-1}.$$

We shall use this formula to determine the value of our integral.

Let $\mathfrak{F}$ denote the space of linear functions on $\mathfrak{h}$. By a rational function ω on $\mathfrak{F}$, we mean an element of the quotient field of the ring of polynomial functions on $\mathfrak{F}$ (see [5(b), p. 194]). We say that ω is defined at a point $\mu \in \mathfrak{F}$, if it can be written in the form $\omega = f/g$ where f and g are two polynomial functions on $\mathfrak{F}$ and $g(\mu) \neq 0$. It is obvious that in that case the ratio $f(\mu)/g(\mu)$ depends only on ω (and not on the choice of f and g). This ratio is called the value of ω at μ and denoted by $\omega(\mu)$. We shall need the following simple lemma on rational functions.

LEMMA 26. *Let f and $g \neq 0$ be two polynomial functions on F and let $\omega = f/g$. Suppose there exists a base $\lambda_1, \cdots, \lambda_l$ for $\mathfrak{F}$ over C and an integer N_0 with the following property. For any given integers $m_1, \cdots, m_l \geqq N_0$, $\omega(m_1\lambda_1 + \cdots + m_l\lambda_l) = 0$ provided $g(m_1\lambda_1 + \cdots + m_l\lambda_l) \neq 0$. Then $f = 0$.*

For otherwise $fg \neq 0$ and therefore we can choose integers $m_1, \cdots, m_l \geqq N_0$ such that $f(m_1\lambda_1 + \cdots + m_l\lambda_l) g(m_1\lambda_1 + \cdots + m_l\lambda_l) \neq 0$ (see Lemma 32 of [5(a)]). Then obviously $\omega(m_1\lambda_1 + \cdots + m_l\lambda_l) \neq 0$ and we get a contradiction.

COROLLARY. *Suppose ω_1, ω_2 are two rational functions on $\mathfrak{F}$ such that*

$$\omega_1(m_1\lambda_1 + \cdots + m_l\lambda_l) = \omega_2(m_1\lambda_1 + \cdots + m_l\lambda_l)$$

for all integral values of $m_1, \cdots, m_l \geqq N_0$ whenever both sides are defined. Then $\omega_1 = \omega_2$.

Let $\omega_1 = f_1/g_1$, $\omega_2 = f_2/g_2$ where f_1, f_2, g_1, g_2 are polynomial functions and $g_1 \neq 0$, $g_2 \neq 0$. Then $\omega_1 - \omega_2 = (g_2 f_1 - g_1 f_2)/g_1 g_2$. Since ω_1 and ω_2 are both defined at any point where $g_1 g_2$ does not vanish, it follows from the above lemma that $g_2 f_1 - g_1 f_2 = 0$ and therefore $\omega_1 = \omega_2$.

For any polynomial function f and $\tau \in \mathfrak{w}_\mathfrak{k}$, we denote by f^τ the polynomial function $\mu \to f(\tau^{-1}\mu)$ $(\mu \in \mathfrak{F})$. Now define a polynomial function F_0 as follows:

$$F_0(\mu) = \prod_{1 \leqq i \leqq s} \prod_{1 \leqq k \leqq s_i} \Big(\sum_{k \leqq j \leqq s_i} q_{ij}(\mu) \Big) \qquad (\mu \in \mathfrak{F})$$

where $q_{ij}(\mu) = \mu(H_{\gamma_p})$ if $\beta_{ij} = \gamma_p$ $(1 \leqq p \leqq s)$ in the notation of Section 8. Similarly let f_0, g_τ and F_τ $(\tau \in \mathfrak{w}_\mathfrak{k})$ be the polynomial functions given by

$$f_0(\mu) = \{w_0 \prod_{\alpha \in C_0} \rho_0(H_\alpha)\} \prod_{\alpha \in P_\mathfrak{k}} \{(\mu(H_\alpha) + \rho(H_\alpha))/\rho(H_\alpha)\}^2$$

$$g(\mu) = \prod_{\alpha \in C_0} \{\mu(\tau H_\alpha) + \rho(\tau H_\alpha)\}$$

$$F_\tau(\mu) = F_0{}^\tau(-2(\mu + \rho)) \qquad (\mu \in \mathfrak{F})$$

and put

$$\omega = 2^p \sum_{\tau \in \mathfrak{w}_\mathfrak{k}} \epsilon(\tau) g_\tau / f_0 F_\tau$$

where $p = 2 \sum_{1 \leqq i \leqq s} \rho_+(H_{\gamma_i})$. Then it follows from what we have proved above that

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\dot{x} = 2^p w c N^{-1} \{f_0(\Lambda)\}^{-1} \int_{\mathfrak{b}_y} \{ \sum_{\tau \in \mathfrak{w}_\mathfrak{k}} \epsilon(\tau) g_\tau(\Lambda) \prod_i y_i^{-2\tau\Lambda'(H_{\gamma_i})-1} \} \, dy.$$

Now let us suppose that $\Lambda'(H_\gamma) < 0$ for every totally positive root γ. Then since any $\tau \in \mathfrak{w}_\mathfrak{k}$ permutes totally positive roots among themselves [5(e), Lemma 10], it follows that $\tau\Lambda'(H_\gamma) < 0$ $(\gamma \in P_+)$. Hence the above-mentioned formula is applicable to each term of the integral on the right and we see that the rational function ω is defined at Λ and

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\dot{x} = w c N^{-1} \omega(\Lambda).$$

Thus we have the following result.

LEMMA 27. *Let $\mathfrak{F}_G(P)$ denote the set of all linear functions on Λ satisfying the following two conditions:*

(1) *Λ is real and $\Lambda(H_\alpha)$ is a non-negative integer for every positive compact root α.*

(2) *$\Lambda(H_\gamma) + \rho(H_\gamma) < 0$ for every totally positive root γ. Then there*

exists a uniquely determined rational function ω *on* $\mathfrak{F}$ *which is defined everywhere on* $\mathfrak{F}_G(P)$ *and such that*

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = wcN^{-1}\omega(\Lambda)$$

for all $\Lambda \in \mathfrak{F}_G(P)$.

We have only to show that ω is unique. Let $(\alpha_1, \cdots, \alpha_l)$ be a fundamental system of positive roots and suppose $(\alpha_1, \cdots, \alpha_m)$ are all the totally positive roots in this system. Define l linear functions Λ_i $(1 \leqq i \leqq l)$ on $\mathfrak{h}$ by the conditions $\Lambda_i(H_{\alpha_j}) = \delta_{ij}$ $1 \leqq i, j \leqq l$. Then if $\lambda_+ = \Lambda_1 + \cdots + \Lambda_m$, it follows from Lemma 13 of [5(e)] that $\lambda_+(H_\gamma) > 0$ for $\gamma \in P_+$ and $\lambda_+(H_\alpha) = 0$ for $\alpha \in P_I$. Hence it is obvious that we can find an integer $n \geqq 2$ such that $\Lambda_i - n\lambda_+ \in \mathfrak{F}_G(P)$ $(1 \leqq i \leqq l)$. Therefore if $\lambda_i = \Lambda_i - n\lambda_+$, $(\lambda_1, \cdots, \lambda_l)$ is a base for $\mathfrak{F}$ over C and $m_1\lambda_1 + \cdots + m_l\lambda_l \in \mathfrak{F}_G(P)$ for every set of positive integers $(m_1, \cdots, m_l)$. The uniqueness of ω now follows immediately from the Corollary to Lemma 26.

We shall now determine ω in another way. Let $\mathfrak{F}_U$ denote the set of all real linear functions Λ on $\mathfrak{h}$ such that $\Lambda(H_\beta)$ is a non-negative integer for every positive root β. For a fixed $\Lambda \in \mathfrak{F}_U$, let σ denote the irreducible complex representation (see [5(f), § 6]) of G_c on a finite-dimensional Hilbert space V with the highest weight Λ. We assume that σ is unitary on U. ϕ_0, being a unit vector in V belonging to the weight Λ, we consider the function

$$\psi_\Lambda^*(x) = (\phi_0, \sigma(x)\phi_0) \qquad (x \in G_c).$$

Since V is irreducible under $\sigma(U)$, it follows from the Schur orthogonality relations for the compact group U, that

$$\int_U |\psi_\Lambda^*(u)|^2 \, du = (\dim V)^{-1}.$$

On the other hand if we normalize the various Haar measures according to Lemma 22 and put $\Lambda' = \Lambda + \rho$, we get

$$\int_U |\psi_\Lambda^*(u)|^2 \, du = w^*N^{-1}\Big(\int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} \, da^*\Big)^{-1}$$

$$\times \int_{B^*} |D(a^*)|^{\frac{1}{2}} \, da^* \int_{K_0 \times K_0} |\psi_\Lambda^*(ka^*k')|^2 \, dk dk'$$

where B^* is defined as in Lemma 24 and we make use of the fact (see the Corollary to Lemma 24) that B^* and B both have N connected components. Now let $H = \sum_i t_i(X_{\gamma_i} + X_{-\gamma_i}) \in \mathfrak{b}^*$ $(t_i \in R)$ and $a^* = e(H)$. Then if we put

$$H^* = \sum_i \log(\cos t_i) H_{\gamma_i}$$

and $h(a^*) = \exp H^* \in A_c$, we know from Lemma 20 that

$$a^* = \zeta h(a^*) \xi$$

where $\zeta \in \mathfrak{P}_c^-$ and $\xi \in \mathfrak{P}_c^+$. Therefore we conclude in the same way as before, that

$$(\phi_0, \sigma(ka^*k')\phi_0) = (\phi_0, \sigma(kh(a^*)k')\phi_0)$$

and therefore

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(ka^*k')\phi_0)|^2 \, dk dk' = \int_{K_0 \times K_0} |(\phi_0, \sigma(kh(a^*)k')\phi_0)|^2 \, dk dk'.$$

On the other hand if V_0 is the subspace of V spanned by $\sigma(k)\phi_0$ $(k \in K_c)$, we prove exactly as before that the corresponding representation σ_0 of K_c on V_0 is irreducible and has the highest weight Λ. Then if χ_Λ denotes the character of σ_0 we have

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(ka^*k')\phi_0)|^2 \, dk dk' = (\dim V_0)^{-2} \chi_\Lambda((h(a^*))^2).$$

Therefore

$$\int_{B^*} |D(a^*)|^{\frac{1}{2}} da^* \int_{K_0 \times K_0} |\psi_\Lambda^*(ka^*k')|^2 \, dk dk'$$

$$= (\dim V_0)^{-2} \int_{B^*} \chi_\Lambda((h(a^*))^2) |D(a^*)|^{\frac{1}{2}} da^*$$

$$= (\dim V_0)^{-2} c \int_{\mathfrak{b}^*} \chi_\Lambda(\exp 2H^*) |D(e(H))|^{\frac{1}{2}} dt$$

where c and dt have the same meaning as before. But from Lemmas 23 and 25 it follows that[10]

$$\chi_\Lambda(\exp 2H^*) |D(e(H))|^{\frac{1}{2}}$$

$$= (-1)^p 2^p \{w_0 \prod_{\alpha \in C_0} \rho(H_\alpha)\}^{-1} \{\prod_{1 \leqq i \leqq s} \sin t_i\}$$

$$\times \sum_{\tau \in \mathfrak{W}_\mathfrak{k}} \{\epsilon(\tau) g_\tau(\Lambda) \prod_{1 \leqq i \leqq s} (\cos t_i)^{2\tau\Lambda'(H_{\gamma_i})-1}\}$$

where $\Lambda' = \Lambda + \rho$. Now put $y_i = \cos t_i$. Then $\mathfrak{b}^*$ corresponds to the region $\mathfrak{b}_y$ defined as before by $0 < y_i < 1$ and $y_i > y_j$ if $\gamma_i \prec \gamma_j$ $(1 \leqq i < j \leqq s)$.

---

[10] The proof of Lemma 25 made no use of the fact that $\Lambda_0(H_\alpha) = 0$ for every totally positive root α lying in the fundamental system $(\alpha_1, \cdots, \alpha_l)$ of positive roots. Hence this lemma is applicable also to χ_Λ.

Therefore if we use Weyl's formula [11(a)] for $\dim V_0$ and recall that $\rho(H_\alpha) = \rho_\mathfrak{k}(H_\alpha)$ $(\alpha \in P_\mathfrak{k})$, we get

$$(\dim V_0)^{-2} \int_{\mathfrak{b}^*} \chi_\Lambda(\exp 2H^*) \mid D(e(H)) \mid^{\frac{1}{2}} dt$$

$$= (-1)^p \{2^p / f_0(\Lambda)\} \int_{\mathfrak{b}_y} \{ \sum_{\tau \in \mathfrak{W}_\mathfrak{k}} \epsilon(\tau) g_\tau(\Lambda) \prod_{1 \leqq i \leqq s} y_i^{2\tau\Lambda'(H_{\gamma_i})-1} \} dy.$$

We have seen that $\tau \in \mathfrak{W}_\mathfrak{k}$ permutes the totally positive roots among themselves and if $\gamma \in P_+$, $\Lambda'(H_\gamma) > 0$ since $\Lambda(H_\gamma) \geqq 0$ and $\rho(H_\gamma) > 0$ (see Weyl [11(b)]). Therefore $2\tau\Lambda'(H_{\gamma_i}) > 0$ and it follows that the right hand side of the above equation is

$$= (-1)^p 2^p \sum_{\tau \in \mathfrak{W}_\mathfrak{k}} \epsilon(\tau) g_\tau(\Lambda) \{ f_0(\Lambda) F_0{}^\tau(2\Lambda') \}^{-1}.$$

Put $\Lambda^* = -(\Lambda + 2\rho)$. Then it is obvious from the definition of F_0, f_0, F_τ and g_τ $(\tau \in \mathfrak{W}_\mathfrak{k})$ that

$$F_\tau(\Lambda^*) = F_0{}^\tau(2\Lambda'), f_0(\Lambda^*) = f_0(\Lambda), g_\tau(\Lambda^*) = (-1)^{r_0} g_\tau(\Lambda)$$

where r_0 is the number of roots in C_0. Hence

$$\int_U \mid \psi_\Lambda{}^*(u) \mid^2 du = w^* c N^{-1} \Big(\int_{\mathfrak{A}^*} \mid D(a^*) \mid^{\frac{1}{2}} da^* \Big)^{-1} (-1)^{p+r_0} \omega(\Lambda^*).$$

But on the other hand by Weyl's formula [11(a)],

$$\int_U \mid \psi_\Lambda{}^*(u) \mid^2 du = (\dim V)^{-1} = \{ \prod_{\beta \in P} \Lambda'(H_\beta)/\rho(H_\beta) \}^{-1}$$

since Λ is the highest weight of σ. This proves that

$$\omega(\Lambda^*) = (-1)^{p+r_0+m} (N/cw^*) \int_{\mathfrak{A}^*} \mid D(a^*) \mid^{\frac{1}{2}} da^*$$
$$\times \{ \prod_{\beta \in P} (\Lambda^*(H_\beta) + \rho(H_\beta))/\rho(H_\beta) \}^{-1}$$

where m is the total number of positive roots. Let $\mathfrak{F}_U{}^*$ denote the set of all linear functions of the form $\Lambda^* = -(\Lambda + 2\rho)$ $(\Lambda \in \mathfrak{F}_U)$. Define the linear functions Λ_i $1 \leqq i \leqq l$ as in the proof of Lemma 27. Then if $\lambda_1 = -(\Lambda_1 + 2\rho)$ and $\lambda_i = -\Lambda_i$ $2 \leqq i \leqq l$, $(\lambda_1, \cdots, \lambda_l)$ is a base for $\mathfrak{F}$ and $m_1\lambda_1 + \cdots + m_l\lambda_l \in \mathfrak{F}_U{}^*$ for positive integers $(m_1, \cdots, m_l)$. Therefore in view of the Corollary to Lemma 26, we can conclude that

$$\omega(\mu) = (-1)^{p+r_0+m} (N/cw^*) \int_{\mathfrak{A}^*} \mid D(a^*) \mid^{\frac{1}{2}} da^*$$
$$\times \{ \prod_{\beta \in P} (\mu(H_\beta) + \rho(H_\beta))/\rho(H_\beta) \}^{-1}$$

($\mu \in \mathfrak{F}$), both sides being defined and equal whenever at least one of them is defined. Hence in particular if $\Lambda \in \mathfrak{F}_G(P)$, it follows from Lemma 27 that

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = (-1)^{p+r_0+m}(w/w^*) \int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^* \times \{\prod_{\beta \in P} \Lambda'(H_\beta)/\rho(H_\beta)\}^{-1}$$

where $\Lambda' = \Lambda + \rho$. On the other hand if q is the number of totally positive roots, $\prod_{\beta \in P} \Lambda'(H_\beta)$ has obviously the same sign as $(-1)^q$. Therefore since the left side is positive and $\rho(H_\beta) > 0$ for every positive root β, $(-1)^{p+r_0+m} = (-1)^q$. Thus we have obtained the following result.

LEMMA 28. *If* $\Lambda \in \mathfrak{F}_G(P)$,

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = (-1)^q (w/w^*) \int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^* \times \{\prod_{\beta \in P} (\Lambda(H_\beta) + \rho(H_\beta))/\rho(H_\beta)\}^{-1}.$$

Although our definition of $\alpha_{\mathfrak{p}_0}$ depends on the order (which we have so far assumed to be fixed) on the space $\mathfrak{F}_R$ of real linear functions on $\mathfrak{h}$, we know from Lemma 22 that the above normalization of the measure $d\bar{x}$ and the numbers w, w^*, $\int_{\mathfrak{A}^*} |D(a^*)|^{\frac{1}{2}} da^*$ are actually independent of this order. Therefore if we normalize $d\bar{x}$ in such a way that

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\bar{x} = |\prod_{\beta \in P} \Lambda(H_\beta) + \rho(H_\beta)/\rho(H_\beta)|^{-1}$$

for all $\Lambda \in \mathfrak{F}_G(P)$, this new normalization is also independent of the particular order in $\mathfrak{F}_R$. However since the definition of the function ψ_Λ depends on this order, we shall now denote it by $\psi_\Lambda{}^P$. Then our result may be stated as follows.

THEOREM 4. *It is possible to normalize the Haar measure on* $d\bar{x}$ *on* G_0 *in such a way that the following condition is fulfilled. Let* P *be the set of all positive roots under any given order* $\mathfrak{F}_R$ *and suppose every noncompact root in* P *is totally positive. Let* $\rho = \frac{1}{2}\sum_{\beta \in P} \beta$ *and let* $\mathfrak{F}_G(P)$ *denote the set of all real linear functions* Λ *on* $\mathfrak{h}$ *which satisfy the following two conditions:*

(1) $\Lambda(H_\alpha)$ *is a non-negative integer for every positive compact root* α.

(2) $\Lambda(H_\beta) + \rho(H_\beta) < 0$ *for noncompact positive root* β.

Then

$$\int_{G_0} |\psi_\Lambda^P(x)|^2\, d\bar{x} = |\prod_{\beta \in P} \Lambda(H_\beta)\, \rho(H_\beta)/\rho(H_\beta)|^{-1}$$

for all $\Lambda \in \mathfrak{F}_G(P)$.

Now we return again to our fixed order in $\mathfrak{F}_R$. Let Λ be a real linear function on $\mathfrak{h}$ such that $\Lambda(H_\alpha)$ is a non-negative integer for every compact positive root α. Define $\psi_\Lambda(x)$ $(x \in G)$ as in the beginning of this section.

LEMMA 29. $\int_{G_0} |\psi_\Lambda(x)|^2\, d\bar{x} < \infty$ *if and only if* $\Lambda \in \mathfrak{F}_G(P)$.

Define λ_+ as in the proof of Lemma 27. Then, as we have seen, $\lambda_+(H_\gamma) > 0$ for $\gamma \in P_+$ and $\lambda_+(H_\alpha) = 0$ if $\alpha \in P_\mathfrak{k}$. Put

$$t_0 = \max_{\gamma \in P_+} \{\Lambda(H_\gamma) + \rho(H_\gamma)\}/\lambda_+(H_\gamma).$$

Then it is obvious that $\Lambda - t\lambda_+ \in \mathfrak{F}_G(P)$ $(t \in R)$ if and only if $t > t_0$ and therefore $t_0 \geqq 0$ if and only if $\Lambda \notin \mathfrak{F}_G(P)$. Moreover if $\Lambda_t = \Lambda - t\lambda_+$, it is clear that

$$\psi_{\Lambda_t}(x) = \psi_\Lambda(x) e^{-t\lambda_+(\Gamma(x))} \qquad (x \in G).$$

But if $H = \sum_{1 \leqq i \leqq s} t_i(X_{\gamma_i} + X_{-\gamma_i})$ $(t_i \in R)$ we have seen above that

$$\lambda_+(\Gamma(\exp H)) = \sum_{1 \leqq i \leqq s} (\log \cosh t_i)\lambda_+(H_{\gamma_i}).$$

Since $\cosh t_i \geqq 1$ and $\lambda_+(H_{\gamma_i}) \geqq 0$, it follows that $\lambda_+(\Gamma(a)) \geqq 0$ $(a \in \mathfrak{A})$. Therefore it is obvious from Lemma 22, that

$$\int_{G_0} |\psi_\Lambda(x)|^2\, d\bar{x} = \int_{G_0} |\psi_{\Lambda_t}(x) e^{t\lambda_+(\Gamma(x))}|^2\, d\bar{x} \geqq \int_{G_0} |\psi_{\Lambda_t}(x)|^2\, d\bar{x}$$

provided $t \geqq 0$. Now suppose $\Lambda \notin \mathfrak{F}_G(P)$ so that $t_0 \geqq 0$. Put $t = t_0 + \epsilon$ where ϵ is positive. Then, in view of Theorem 4, we get

$$\int_{G_0} \psi_\Lambda(x)|^2\, d\bar{x} \geqq |\prod_{\beta \in P} \{\Lambda(H_\beta) + \rho(H_\beta) - (t_0 + \epsilon)\lambda_+(H_\beta)\}/\rho(H_\beta)|^{-1}.$$

But it follows from the definition of t_0 that as $\epsilon \to 0$ the right side tends to infinity. Therefore

$$\int_{G_0} |\psi_\Lambda(x)|^2\, d\bar{x} = \infty.$$

Conversely if $\Lambda \in \mathfrak{F}_G(P)$ we know from Theorem 4 that

$$\int_{G_0} |\psi_\Lambda(x)|^2\, d\bar{x} < \infty.$$

Thus the lemma is proved.

We shall now consider the question of the integrability of the function ψ_Λ. Let us recall that $\rho_+ = \frac{1}{2}\sum_{\beta \in P_+} \beta$.

LEMMA 30. *Let Λ be a linear function on $\mathfrak{h}$ satisfying the following two conditions:*

(1) *$\Lambda(H_\alpha)$ is a non-negative integer for every compact positive root α.*

(2) *$\Lambda(H_\beta) + \rho(H_\beta) < 1 - 2\rho_+(H_\beta)$ for every noncompact positive root β.*

Then

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\dot{x} < \infty.$$

We use the notation which was introduced at the beginning of this section. In view of Lemma 22, it is enough to prove that

$$\int_B |D(a)|^{\frac{1}{2}} da \int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)| \, e^{\lambda(\Gamma(a))} \, dkdk' < \infty.$$

But it follows from the Schwartz inequality that

$$\begin{aligned}\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)| \, dkdk' &\leqq \{\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)|^2 \, dkdk'\}^{\frac{1}{2}} \\ &= (\dim V_0)^{-1}\{\mathrm{Sp}\,\sigma_0((h(a))^2)\}^{\frac{1}{2}}\end{aligned}$$

as we have seen before. Since $\sigma_0(h(a))$ is obviously a positive definite self-adjoint transformation on V_0, it is clear that

$$\mathrm{Sp}\,\sigma_0((h(a))^2) \leqq \{\mathrm{Sp}\,\sigma_0(h(a))\}^2.$$

Therefore

$$\int_{K_0 \times K_0} |(\phi_0, \sigma(kak')\phi_0)| \, dkdk' \leqq (\dim V_0)^{-1} \chi_{\Lambda_0}(h(a)).$$

So it would be sufficient to prove that

$$\int_B \chi_{\Lambda_0}(h(a)) \, e^{\lambda(\Gamma(a))} \, |D(a)|^{\frac{1}{2}} da < \infty.$$

For any point

$$H = \sum_{1 \leqq i \leqq s} t_i(X_{\gamma_i} + X_{-\gamma_i}) \quad (t \in R) \text{ in } \mathfrak{b} \text{ we put } H' = \sum_{1 \leqq i \leqq s} (\log \cosh t_i) H_{\gamma_i}$$

as before. Then if $a = \exp H$, $h(a) = \exp H'$ and therefore it follows from Lemmas 23 and 25 that

$$\begin{aligned}&\chi_{\Lambda_0}(h(a)) \, e^{\lambda(\Gamma(a))} \, |D(a)|^{\frac{1}{2}} \\ &= c_1 \frac{\Delta(2H')}{\Delta(H')} \{\sum_{\tau \in \mathfrak{W}_f} \epsilon(\tau) g_\tau(\Lambda) \prod_{i=1}^{s} (\cosh t_i)^{\tau\Lambda'(H_{\gamma_i}) + \rho_+(H_{\gamma_i}) - 1}\} \prod_{i=1}^{s} (\sinh t_i)\end{aligned}$$

where $\Lambda' = \Lambda + \rho$, c_1 is a positive constant and g_τ is defined as before. Now

$$\frac{\Delta(2H')}{\Delta(H')} = \prod_{\alpha \in C'} (e^{\frac{1}{2}\alpha(H')} + e^{-\frac{1}{2}\alpha(H')})$$

and therefore if r' is the number of roots in C', it follows from Lemma 16 that

$$0 < \frac{\Delta(2H')}{\Delta(H')} \leqq 2^{r'} \prod_{\alpha \in C'} e^{\frac{1}{2}|\alpha(H')|}$$

$$= 2^{r'} \prod_{i=1}^{s} (\cosh t_i)^{\frac{1}{2}r_i} \prod_{1 \leqq i < j \leqq s} (\cosh t_j / \cosh t_i)^{\frac{1}{2}r_{ij}}$$

in the notation of Section 6. (Here we have to make use of the fact that $t_j > t_i$ on $\mathfrak{b}$ if $r_{ij} > 0$ $(1 \leqq i < j \leqq s)$). But since

$$r_i - \sum_{i<j\leqq s} r_{ij} + \sum_{1\leqq j<i} r_{ji} \leqq 2\rho_+(H_{\gamma_i}) - 2$$

from Lemma 18, we conclude that

$$0 < \{\Delta(2H')/\Delta(H')\} \leqq 2^{r'} \prod_{i=1}^{s} (\cosh t_i)^{\rho_+(H_{\gamma_i})-1}$$

and therefore if $c_2 = 2^{r'} c_1$,

$$\chi_{\Lambda_0}(h(a))\, e^{\lambda(\Gamma(a))}\, |D(a)|^{\frac{1}{2}}$$

$$\leqq c_2 \{ \sum_{\tau \in \mathfrak{W}_\mathfrak{l}} \epsilon(\tau) g_\tau(\Lambda) \prod_i (\cosh t_i)^{\tau\Lambda'(H_{\gamma_i})+2\rho_+(H_{\gamma_i})-2} \} \prod_i \sinh t_i.$$

Now put $y_i = (\cosh t_i)^{-1}$, $dy = dy_1 \cdots dy_s$ and define $\mathfrak{b}_y$ as before. Then

$$\int_{\mathfrak{b}} \prod_i (\cosh t_i)^{\tau\Lambda'(H_{\gamma_i})+2\rho_+(H_{\gamma_i})-2} \prod_i (\sinh t_i)\, dt$$

$$= \int_{\mathfrak{b}_y} \prod_i y_i^{-\tau\Lambda'(H_{\gamma_i})-2\rho_+(H_{\gamma_i})}\, dy \qquad (\tau \in \mathfrak{W}_\mathfrak{l}).$$

But we have seen earlier that $\tau\rho_+ = \rho_+$ and therefore

$$\tau\Lambda'(H_{\gamma_i}) + 2\rho_+(H_{\gamma_i}) = \Lambda'(H_\gamma) + 2\rho_+(H_\gamma)$$

where $\gamma = \tau^{-1}\gamma_i$. Hence it follows from our hypothesis on Λ that

$$-\tau\Lambda'(H_{\gamma_i}) - 2\rho_+(H_{\gamma_i}) + 1 > 0$$

and therefore

$$\int_{\mathfrak{b}_y} \prod_i y_i^{-\tau\Lambda'(H_{\gamma_i})-2\rho_+(H_{\gamma_i})}\, dy < \infty.$$

This shows that

$$\int_B \chi_{\Lambda_0}(h(a))\, e^{\lambda(\Gamma(a))}\, |D(a)|^{\frac{1}{2}}\, da < \infty$$

and so the lemma is proved.

10. Integrable and square-integrable representations. Suppose Λ is a real linear function on $\mathfrak{h}$ which satisfies the first condition of Lemma 30. We define the Hilbert space $\mathfrak{H}_\Lambda$ as in Theorem 2 of [5(f)] corresponding to the function $\mu = 1$ on G_0. We know from Lemma 14 of [5(f)] and Lemma 29 that $\mathfrak{H}_\Lambda \neq 0$ if and only if $\Lambda \in \mathfrak{F}_G(P)$. So now let us assume that $\Lambda \in \mathfrak{F}_G(P)$ and let π_Λ denote the representation of G on $\mathfrak{H}_\Lambda$ (see [5(f), Theorem 2]). Then π_Λ is irreducible and unitary and

$$(\psi_\Lambda, \pi_\Lambda(x)) = \psi_\Lambda(x) \| \psi_\Lambda \|^2 \qquad (x \in G)$$

from the Corollary to Theorem 2 of [5(f)]. (Here we use the usual notation for the norm and the scalar product in $\mathfrak{H}_\Lambda$). Since $\psi_\Lambda \neq 0$ and

$$\int_{G_0} |\psi_\Lambda(x)|^2 \, d\dot{x} = \| \psi_\Lambda \|^2 < \infty,$$

this prove that π_Λ is square-integrable (see Section 3). Moreover if Λ satisfies the second condition of Lemma 30, we prove in the same way that π_Λ is integrable. Combining these results with Lemma 19 of [5(e)], we get Theorem 3 of [5(e)] which was stated there without proof.

It is clear from the definition of $\mathfrak{H}_\Lambda$ that every element $\phi \in \mathfrak{H}_\Lambda$ is compeletly determined by its restriction on G and in fact

$$\| \phi \|^2 = \int_{G_0} |\phi(x)|^2 \, d\dot{x}.$$

Therefore since $\psi_\Lambda \in \mathfrak{H}_\Lambda$ and π_Λ is irreducible, $\mathfrak{H}_\Lambda$ may be identified with the completion (under the above norm) of the space spanned by the right translates of the function ψ_Λ under G. If we use the normalization of Theorem 4 for the Haar measure on G_0 and denote by d_Λ the formal degree of π_Λ (corresponding to the kernel Z_0 of the mapping $x \to \dot{x}$ of G on G_0 (see Section 3)), it follows from Theorem 1 that

$$d_\Lambda^{-1} \| \psi_\Lambda \|^2 = \int_{G_0} |(\psi_\Lambda, \pi_\Lambda(x)\psi_\Lambda)|^2 \, d\dot{x} = \| \psi_\Lambda \|^4$$

and therefore

$$d_\Lambda = \| \psi_\Lambda \|^{-2} = \prod_{\beta \in P} |(\Lambda(H_\beta) + \rho(H_\beta))/\rho(H_\beta)|$$

from Theorem 4. This is the formula for the formal degree in terms of the highest weight Λ of the representation. It is substantially the same as the corresponding formula of Weyl [11(a)] for a compact semisimple group.

11. Similarity with finite-dimensional representations. In order to simplify matters let us suppose in this section that G is simple but not compact. Then if $(\alpha_0, \alpha_1, \cdots, \alpha_l)$ is a fundamental system of positive roots, we know (Corollary 2 to Lemma 13 of [5(e)]) that it contains exactly one noncompact root, which we may assume to be α_0. Let Λ_i $0 \leqq i \leqq l$ be the linear functions on $\mathfrak{h}$ given by $\Lambda_i(H_{\alpha_j}) = \delta_{ij}$ $0 \leqq i, j \leqq l$ and let σ_i be an irreducible complex representation of G_c (see [5(f), § 6]) on a finite-dimensional Hilbert space V_i with the highest weight Λ_i. We assume that σ_i is unitary on V. Let ϕ_i be a unit vector in V_i belonging to the weight Λ_i. Then if $\psi_i(z) = (\phi_i, \sigma_i(z)\phi_i)$ $(z \in G_c)$, it follows from Lemmas 6 and 14 of [5(f)] that $\psi_i(\bar{x}) = \psi_{\Lambda_i}(x)$ $(0 \leqq i \leqq l)$ and

$$\psi_0(\bar{x}) = \psi_{\Lambda_0}(x) = \exp(\Lambda_0(\Gamma(x))) \qquad (x \in G).$$

Hence in particular $\psi_{\Lambda_0}(x)$ is never zero. Now if Λ is any linear function on $\mathfrak{h}$, it is obvious that $\Lambda = \lambda_0\Lambda_0 + \cdots + \lambda_l\Lambda_l$ where $\lambda_i = \Lambda(H_{\alpha_i})$ $0 \leqq i \leqq l$. Therefore if $\Lambda \in \mathfrak{F}_G(P)$, $\lambda_1, \cdots, \lambda_l$ are all non-negative integers while λ_0 is a *negative* real number. Moreover it follows again from Lemma 6 of [5(f)] that

$$\psi_\Lambda(x) = e^{\lambda_0\Lambda_0(\Gamma(x))}\psi_1^{\lambda_1}(\bar{x}) \cdots \psi_l^{\lambda_l}(\bar{x}) \qquad (x \in G).$$

In particular if λ_0 is also an integer, $\psi_\Lambda(x)$ depends only on $\bar{x}$ and so if we regard ψ_Λ as a function on G_0, we have

$$\psi_\Lambda = \psi_0^{\lambda_0}\psi_1^{\lambda_1} \cdots \psi_l^{\lambda_l}.$$

On the other hand if $m_0, m_1, \cdots, m_l$ are non-negative integers,

$$\psi = \psi_0^{m_0}\psi_1^{m_1} \cdots \psi_l^{m_l}$$

is a holomorphic function on G_c and again we can conclude from Lemma 6 of [5(f)] that

$$\psi(z) = (\phi, \sigma(z)\phi) \qquad (z \in G_c)$$

where σ is the irreducible complex representation of G_c on a finite-dimensional Hilbert space V with the highest weight $m_0\Lambda_0 + \cdots + m_l\Lambda_l$ and ϕ is a unit vector in V belonging to this weight. (We assume of course that σ is unitary on U). Then $\psi(u)$ $(u \in U)$ is a matrix coefficient of an irreducible representation of U and therefore the space spanned by all the right translates of ψ under U is irreducible under the corresponding representation of U. Thus the similarity between this case and that of $\psi_\cdot$, $\mathfrak{H}_\Lambda$ and π_Λ discussed above is now quite obvious.

12. Proof of Lemma 22. In order to prove Lemma 22, we shall use a method which has been extensively used before by Weyl [11(a)] and Cartan [2(a)]. It is no longer necessary to assume that $\mathfrak{h}_0$ is maximal abelian in $\mathfrak{g}_0$, since this assumption plays no role whatever in our proof. However we still consider $\mathfrak{g}$ as a Hilbert space under the norm $\| X \|^2 = -B(\theta(X), X)$ $(X \in \mathfrak{g})$ and denote by G_c a simply connected complex Lie group with the Lie algebra $\mathfrak{g}$. As before G_0, K_0 and U are the (real) analytic subgroups of G_c corresponding to $\mathfrak{g}_0$, $\mathfrak{k}_0$ and $\mathfrak{u} = \mathfrak{k}_0 + (-1)^{\frac{1}{2}}\mathfrak{p}_0$. We assume that $\int_{K_0} dk = 1$.

Let us introduce the following notation for the sake of convenience. Suppose μ and μ' are positive measures on two locally compact spaces E and E' respectively and f is an open continuous mapping of E into E' which is locally one-one on E. Then we write $d\mu \underset{f}{\sim} d\mu'$ if there exists a positive constant c with the following property. Let U be an open set in E such that f is univalent on U. Then $\mu(U) = c\mu'(f(U))$. If it is clear from the context which mapping f we have in mind, we write simply $d\mu \sim d\mu'$.

LEMMA[11] 31. *$(k, X) \to k \exp X$ $(k \in K_0, X \in \mathfrak{p}_0)$ is a one-one regular analytic mapping of $K_0 \times \mathfrak{p}_0$ onto G_0.*

Let ϕ denote this mapping. It is known (Cartan [2(b)], Mostow [9]) that ϕ is one-one and onto G_0. Also it is obviously analytic. Let f be a function on G_0 which is defined and analytic around $x = k \exp X$. Then if $Y \in \mathfrak{p}_0$ and $Z \in \mathfrak{k}_0$,

$$\left\{ \frac{d}{dt} f(k \exp(X + tY)) \right\}_{t=0} = (Y'f)(x)$$

$$\left\{ \frac{d}{dt} f(k \exp tZ \exp X) \right\}_{t=0} = (Z'f)(x) \qquad (t \in R)$$

where (see Chevalley [3, p. 157])

$$Y' = \{(1 - \exp(-adX)))/adX\}Y$$

$$Z' = \exp(-adX)Z$$

and $(1 - \exp(-adX))/adX$ stands for the sum of the convergent series

$$\sum_{m \geqq 0}^{\infty} (-1)^m (adX)^m/(m+1)!.$$

[11] The proofs of Lemmas 31 and 33, which we present here, are substantially the same as those of Cartan. However in view of later applications, we are interested not only in verifying that certain determinants are different from zero but also in computing their actual value. This compels us to reproduce the whole argument.

Let T be the linear transformation of $\mathfrak{g}$ such that $TY = Y'$ and $TZ = Z'$. Since $(adX)\mathfrak{p} \subset \mathfrak{k}$ and $(adX)^2\mathfrak{p} \subset \mathfrak{p}$, it follows that

$$\exp(adX)Z' = Z$$

$$\exp(adX)Y' \equiv \{\sinh adX/adX\}Y \bmod k$$

where $(\sinh adX/adX) = s(X)$ is defined by the power series

$$\sum_{m \geqq 0} (adX)^{2m}/(2m+1)!.$$

Since only even powers of adX appear in this series $s(X)$ leaves $\mathfrak{p}$ invariant. Also $\tilde{\theta}(X) = -X$ and therefore adX is self-adjoint. Hence $s(X)$ is a positive definite self-adjoint transformation. Therefore the same holds for its restriction $(s(X))_{\mathfrak{p}}$ on $\mathfrak{p}$. On the other hand it is obvious from the above congruence that

$$\det(\exp(adX)T) = \det((s(X))_{\mathfrak{p}}).$$

Since G_c is semisimple $\det(\exp(adX)) = 1$ and therefore

$$\det T = \det(s(X))_{\mathfrak{p}}.$$

The right side is positive since $(s(X))_{\mathfrak{p}}$ is positive definite. This proves that $\det T \neq 0$ and therefore ϕ is regular at (k, X).

Let dX denote the element of volume in p_0 corresponding to the Euclidean metric $\| X \|$ $(X \in \mathfrak{p}_0)$.

COROLLARY. $dx \sim \det(\sinh adX/adX)_{\mathfrak{p}}\, dk dX$ $(x = k \exp X)$.

This follows immediately from our calculation above.

We now state a lemma which will be needed frequently.

LEMMA 32. *Let M and N be two manifolds of class C^1 satisfying the countability axioms. Then if f is a differentiable mapping of M into N, and Q is a subset of M then $\partial\mathrm{im}\, f(Q) \leqq \partial\mathrm{im}\, Q$.*

Here we use the Brouwer-Urysohn-Menger definition[12] of dimension (Hurewicz and Wallman [6]). Although this result must be regarded as known, no easily accessible published proof of it seems to be available. Therefor we shall give a short proof in the Appendix (Section 13).

Let $\mathfrak{a}_{\mathfrak{p}_0}$ be a maximal abelian subspace of $\mathfrak{p}_0$. Then we have the following result[11] due to Cartan [2(a), p. 354].

---

[12] In order to distinguish it from the vector dimension of a complex vector space, we denote the topological dimension by "∂im."

LEMMA 33. $\mathfrak{p}_0 = \bigcup_{k \in K_0} \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$.

Extend $\mathfrak{a}_{\mathfrak{p}_0}$ to a maximal abelian subalgebra $\mathfrak{a}_0$ of $\mathfrak{g}_0$ and let $\mathfrak{a}_\mathfrak{p}$ and $\mathfrak{a}$ be the complexifications of $\mathfrak{a}_{\mathfrak{p}_0}$ and $\mathfrak{a}_0$ respectively in $\mathfrak{g}$. Then $\mathfrak{a}$ is a Cartan subalgebra of $\mathfrak{g}$. We consider the set Σ of all roots of $\mathfrak{g}$ (with respect to $\mathfrak{a}$) which do not vanish identically on $\mathfrak{a}_{\mathfrak{p}_0}$. For each $\alpha \in \Sigma$ let σ_α denote the hyperplane consisting of all points $H \in \mathfrak{a}_{\mathfrak{p}_0}$ such that $\alpha(H) = 0$ and let $\mathfrak{a}'_{\mathfrak{p}_0}$ denote the complement of $\bigcup_{\alpha \in \Sigma} \sigma_\alpha$ in $\mathfrak{a}_{\mathfrak{p}_0}$. Moreover let M be the set of all $k \in K_0$ such that $\mathrm{Ad}(k)H = H$ for every $H \in \mathfrak{a}_{\mathfrak{p}_0}$. Obviously M is a closed subgroup of K_0. Let $k \to \bar{k}$ $(k \in K_0)$ denote the natural mapping of K_0 on the factor space $\bar{K}_0 = K_0/M$ consisting of the cosets of the form kM. Consider the mapping ϕ of $\bar{K}_0 \times \mathfrak{a}_{\mathfrak{p}_0}$ into $\mathfrak{p}_0$ given by $\phi(\bar{k}, H) = H^k = \mathrm{Ad}(k)H$ where k is any element in the coset $\bar{k}$. ϕ is obviously analytic. Let $\mathfrak{m}_0$ be the centralizer of $\mathfrak{a}_{\mathfrak{p}_0}$ in $\mathfrak{k}_0$, so that $\mathfrak{m}_0$ is the Lie algebra of M. Let f be a function on $\mathfrak{p}_0$ defined and analytic around $X = H_0^{k_0}$ $(H_0 \in \mathfrak{a}_{\mathfrak{p}_0}, k_0 \in K_0)$. Then if $Z \in \mathfrak{k}_0$ and $H \in \mathfrak{a}_{\mathfrak{p}_0}$,

$$\left\{ \frac{d}{dt} f(\phi(\overline{k_0 \exp tZ}, H_0)) \right\}_{t=0} = \left\{ \frac{d}{dt} f(X - t\,\mathrm{Ad}(k_0)[H_0, Z]) \right\}_{t=0}$$

$$\left\{ \frac{d}{dt} f(\phi(\bar{k}_0, H_0 + tH)) \right\}_{t=0} = \left\{ \frac{d}{dt} f(X + t\,\mathrm{Ad}(k_0)H) \right\}_{t=0}$$

$(t \in R)$. Now if $H_0 \in \mathfrak{a}'_{\mathfrak{p}_0}$, $\dim([H_0, \mathfrak{k}]) = \dim \mathfrak{k} - \dim \mathfrak{m}$ (where $\mathfrak{m}$ is the complexification of $\mathfrak{m}_0$). Also $\mathfrak{a}_\mathfrak{p}$ and $(\mathrm{ad}\, H_0)\mathfrak{k}$ are mutually orthogonal. Hence

$$\dim(\mathfrak{a}_\mathfrak{p} + (\mathrm{ad}\, H_0)\mathfrak{k}) = \dim \mathfrak{a}_\mathfrak{p} + \dim \mathfrak{k} - \dim \mathfrak{m} = \dim \mathfrak{p}$$

from Lemma 4 of [5(b)]. This proves that

$$\mathrm{Ad}(k_0)(\mathfrak{a}_\mathfrak{p} + (\mathrm{ad}\, H_0)\mathfrak{k}) = \mathfrak{p}.$$

Since $\dim(\bar{K}_0 \times \mathfrak{a}_{\mathfrak{p}_0}) = \dim \mathfrak{k} - \dim \mathfrak{m} + \dim \mathfrak{a}_\mathfrak{p} = \dim \mathfrak{p} = \dim_R \mathfrak{p}_0$, it follows that ϕ is regular on $\bar{K}_0 \times \mathfrak{a}'_{\mathfrak{p}_0}$. Let $d\bar{k}$ denote the invariant measure on $\bar{K}_0$ such that $\int_{\bar{K}_0} d\bar{k} = 1$. Similarly let dH denote the element of volume in $\mathfrak{a}_{\mathfrak{p}_0}$ corresponding to the Euclidean metric $\| H \|$. Now introduce some (fixed) lexicographic order among roots and let Σ^+ be the set of positive roots in Σ under this order. For each root α select an element $X_\alpha \neq 0$ in $\mathfrak{g}$ such that $[H, X_\alpha] = \alpha(H)X_\alpha$ for all $H \in \mathfrak{a}$. Let $\mathfrak{n}^+$ and $\mathfrak{n}^-$ respectively be the subspaces of $\mathfrak{g}$ spanned by X_α and $X_{-\alpha}$ $(\alpha \in \Sigma^+)$. Then (see the proof of Lemma 4 of [5(b)])

$$\mathfrak{g} = \mathfrak{n}^- + \mathfrak{a}_\mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+ = \mathfrak{k} + \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^+ = \mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+$$

where the sums are all direct. Hence

$$\mathfrak{k}/\mathfrak{m} \cong \mathfrak{g}/(\mathfrak{m} + \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^+) \cong \mathfrak{p}/\mathfrak{a}_\mathfrak{p}$$

and

$$\mathfrak{g}/(\mathfrak{m} + \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^+) \cong \mathfrak{n}^-.$$

Therefore we may identify $\mathfrak{k}/\mathfrak{m}$ and $\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$ with $\mathfrak{n}^-$ in a natural way. Now if $H \in \mathfrak{a}_\mathfrak{p}, \mathfrak{m}$ is contained in the kernel of the mapping $Z \to [H, Z]$ $(Z \in \mathfrak{k})$ and therefore by going to the residue classes we get a mapping of $\mathfrak{k}/\mathfrak{m}$ into $\mathfrak{p}/\mathfrak{a}_\mathfrak{p}$. In view of the above identification, this gives a linear transformation T_H in $\mathfrak{n}^-$. It is clear that T_H coincides with the restriction of $\operatorname{ad} H$ on $\mathfrak{n}^-$. Therefore in view of our calculation above we can conclude that

$$dX \sim |\det T_H| \, dH d\bar{k} = \prod_{\alpha \in \Sigma^+} |\alpha(H)| \, dH d\bar{k} \qquad (X = H^{\bar{k}})$$

$(H \in \mathfrak{a}'_{\mathfrak{p}_0}, \bar{k} \in \bar{K}_0)$. We shall need this result a little later.

Let α be a root in Σ. Since α takes real values on $\mathfrak{a}_{\mathfrak{p}_0}$, we can find an element $X \neq 0$ in $\mathfrak{g}_0$ such that $[H, X] = \alpha(H)X$ for all $H \in \mathfrak{a}_{\mathfrak{p}_0}$. We claim $X \notin \mathfrak{p}_0$. For otherwise since $[\mathfrak{p}_0, \mathfrak{p}_0] \subset \mathfrak{k}_0$, we get $[H, X] \in \mathfrak{k} \cap \mathfrak{p}_0 = 0$. This however is impossible since $X \neq 0$ and $\alpha(H) \neq 0$ for a suitable H in $\mathfrak{a}_{\mathfrak{p}_0}$. Similarly we prove that $X \notin \mathfrak{k}_0$. Therefore $X = Y + Z$ $(Y \in \mathfrak{p}, Z \in \mathfrak{k}_0)$ and $Y \neq 0$, $Z \neq 0$. Then comparing components in $\mathfrak{k}_0$ and $\mathfrak{p}_0$ of both sides of the equation $[H, X] = \alpha(H)X$, we find that $[H, Y] = \alpha(H)Z$ and $[H, Z] = \alpha(H)Y$. Thus we have obtained the following result (see Cartan [2(a), p. 357]).

LEMMA 34. *For each $\alpha \in \Sigma$ we can choose nonzero elements Y_α, Z_α in $\mathfrak{p}_0$ and $\mathfrak{k}_0$ respectively such that*

$$[H, Y_\alpha] = \alpha(H)Z_\alpha, \qquad [H, Z_\alpha] = \alpha(H)Y_\alpha$$

for all $H \in \mathfrak{a}_\mathfrak{p}$.

Let M_α $(\alpha \in \Sigma^+)$ denote the set of all elements $k \in K_0$ such that $\operatorname{Ad}(k)H = H$ for every $H \in \sigma_\alpha$. Then M_α is a compact subgroup of K_0 containing M. Moreover it is obvious that the element Z_α of the above lemma does not lie in $\mathfrak{m}_0$ while it is certainly contained in the Lie algebra of M_α. Hence $\dim M_\alpha > \dim M$. Let ϕ_α be the mapping of $(K_0/M_\alpha) \times \sigma_\alpha$ into $\mathfrak{p}_0$ defined by $\phi_\alpha(k^*, H) = \operatorname{Ad}(k)H$ where $k \in K_0$, $H \in \sigma_\alpha$ and $k^* = kM_\alpha$.

It is obvious that ϕ_α is an analytic mapping and therefore it follows from Lemma 32 that

$$\partial\dim \phi_\alpha((K_0/M_\alpha)\times\sigma_\alpha) \leqq \partial\dim((K_0/M_\alpha)\times\sigma_\alpha) \leqq \dim_R \mathfrak{p}_0 - 2$$

since $\dim(K_0/M_\alpha) < \dim(K_0/M)$ and

$$\partial\dim\sigma_\alpha = \dim_R \sigma_\alpha = \dim_R \mathfrak{a}_{\mathfrak{p}_0} - 1.$$

On the other hand it is clear that

$$\phi_\alpha((K_0/M_\alpha)\times\sigma_\alpha) = \phi(\bar{K}_0\times\sigma_\alpha)$$

and therefore

$$\phi((K_0/M)\times(\bigcup_{\alpha\in\Sigma^+}\sigma_\alpha)) = \bigcup_{\alpha\in\Sigma^+}\phi_\alpha((K_0/M_\alpha)\times\sigma_\alpha).$$

Hence if we denote this set by $\mathfrak{p}_s$,

$$\partial\dim\mathfrak{p}_s \leqq \partial\dim\mathfrak{p}_0 - 2.$$

Therefore the complement ${}^c\mathfrak{p}_s$ of $\mathfrak{p}_s$ in $\mathfrak{p}_0$ is connected (see Hurewicz and Wallman [6, p. 48, Theorem IV 4]). On the other hand let

$$\mathfrak{p}'_0 = \phi(\bar{K}_0\times\mathfrak{a}'_{\mathfrak{p}_0}).$$

Since ϕ is regular and therefore open on $\bar{K}_0\times\mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{p}'_0$ is open in $\mathfrak{p}_0$. Let $r_0 = \dim_R(\mathfrak{a}_{\mathfrak{p}_0}+\mathfrak{m}_0)$ and for any $X\in\mathfrak{p}_0$ let $r(X)$ denote the complex dimension of the centralizer of X in $\mathfrak{g}$. Then it is obvious that $r(X) = r_0$ if $X\in\mathfrak{p}'_0$ while $r(X) > r_0$ if $X\in\mathfrak{p}_s$. Hence $\mathfrak{p}'\subset{}^c\mathfrak{p}_s$. Finally since $\bar{K}_0$ is compact, it follows immediately that $\mathfrak{p}'_0$ is closed in ${}^c\mathfrak{p}_s$. Therefore since ${}^c\mathfrak{p}_s$ is connected and $\mathfrak{p}'_0$ is not empty (because $\mathfrak{a}'_{\mathfrak{p}_0}\subset\mathfrak{p}'_0$) we conclude that $\mathfrak{p}'_0 = {}^c\mathfrak{p}_s$. Now let λ be an indeterminate. Then it is clear that there exists a polynomial function F on $\mathfrak{p}$ such that if $X\in\mathfrak{p}$, $F(X)$ is the coefficient of λ^{r_0} in the characteristic polynomial of $\operatorname{ad}X$ (in λ). It is obvious that $r(X) > r_0$ if and only if $F(X) = 0$. Hence F is not identically zero and $\mathfrak{p}_s$ is contained in the set of zeros of F on $\mathfrak{p}_0$. This shows that ${}^c\mathfrak{p}_s = \mathfrak{p}'_0$ is everywhere dense in $\mathfrak{p}_0$. But $\bar{K}_0$ is compact, and $\|H^k\| = \|H\|$ ($k\in\bar{K}$, $H\in\mathfrak{a}_{\mathfrak{p}_0}$) and so it follows that $\phi(\bar{K}_0\times\mathfrak{a}_{\mathfrak{p}_0})$ is closed in $\mathfrak{p}_0$. Therefore $\phi(\bar{K}_0\times\mathfrak{a}_{\mathfrak{p}_0}) = \mathfrak{p}_0$. This completes the proof of Lemma 33.

Let M' be the normalizer of $\mathfrak{a}_{\mathfrak{p}_0}$ in K. Then M' is a closed subgroup of K_0 and M is a normal subgroup of M'. If $H\in\mathfrak{a}_{\mathfrak{p}_0}$, $\operatorname{ad}H$ is self-adjoint and therefore $(\operatorname{ad}H)^2X = 0$ implies $(\operatorname{ad}H)X = 0$ ($X\in\mathfrak{g}$). From this it follows immediately that the Lie algebra of M' is also $\mathfrak{m}_0$ and therefore $W = M'/M$

is discrete. But since it is also compact, W must be finite. It operates as a group of linear transformations on $\mathfrak{a}_{\mathfrak{p}_0}$ as follows:

$$sH = \mathrm{Ad}(k)H \qquad (s \in W)$$

where k is any element in M' belonging to the coset s.

LEMMA 35. *Let H be an element in $\mathfrak{a}'_{\mathfrak{p}_0}$. Then if $k \in K_0$ and $\mathrm{Ad}(k)H \in \mathfrak{a}_{\mathfrak{p}_0}$, k must lie in M'.*

Since $H \in \mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{a}_{\mathfrak{p}_0}$ is exactly the centralizer of H in $\mathfrak{p}_0$. Therefore $\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ is the centralizer of $\mathrm{Ad}(k)H$ in $\mathfrak{p}_0$. But since $\mathrm{Ad}(k)H \in \mathfrak{a}_{\mathfrak{p}_0}$ and $\mathfrak{a}_{\mathfrak{p}_0}$ is abelian, $\mathfrak{a}_{\mathfrak{p}_0} \subset \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$. However $\mathfrak{a}_{\mathfrak{p}_0}$ and $\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ obviously have the same dimension over R. Therefore $\mathfrak{a}_{\mathfrak{p}_0} = \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$ and so $k \in M'$.

W also operates on $\bar{K}_0$ on the right as follows. Let $\bar{k} \in \bar{K}_0$ and $s \in W$. Since M is a normal subgroup of M', $\bar{k}s = km'M$ where k and m' are any two elements of K_0 and M' respectively lying in the cosets $\bar{k}$ and s. It is obvious that $\phi(\bar{k}s, s^{-1}H) = \phi(\bar{k}, H)$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$. Conversely if $H \in \mathfrak{a}'_{\mathfrak{p}_0}$ and $\phi(\bar{k}, H) = \phi(\bar{k}_1, H_1)$ $(k, k_1 \in K_0, H_1 \in \mathfrak{a}_{\mathfrak{p}_0})$, it follows from Lemma 35 that $k_1^{-1}k \in M'$. This shows that $\bar{k}_1 = \bar{k}s$ and $H_1 = s^{-1}H$ for some $s \in W$. Then if w_0 is the order of the group W, it follows that every element in $\mathfrak{p}'_0$ has exactly w_0 distinct pre-images in $\bar{K}_0 \times \mathfrak{a}'_{\mathfrak{p}_0}$ under ϕ. We have seen above that there exists a polynomial function $F \not\equiv 0$ on $\mathfrak{p}_0$ such that $F(X) = 0$ $(X \in \mathfrak{p}_0)$ if and only if $X \notin \mathfrak{p}'_0$. This shows that the complement of $\mathfrak{p}'_0$ in $\mathfrak{p}_0$ is of Euclidean measure zero. Similarly it is obvious that the complement of $\mathfrak{a}'_{\mathfrak{p}_0}$ in $\mathfrak{a}_{\mathfrak{p}}$ is of Euclidean measure zero. Moreover we have seen above that

$$dX \sim \prod_{\alpha \in \Sigma^+} |\alpha(H)|\, dH d\bar{k} \quad (X = H^{\bar{k}}, H \in \mathfrak{a}'_{\mathfrak{p}_0}, \bar{k} \in \bar{K}_0),$$

and so we have the following result.

LEMMA 36. *There exists a positive constant c with the following property. If f is any continuous function on $\mathfrak{p}_0$ vanishing outside a compact set,*

$$\int_{\mathfrak{p}_0} f(X)\, dX = c \int_{\bar{K}_0 \times \mathfrak{a}_{\mathfrak{p}_0}} f(H^{\bar{k}}) \prod_{\alpha \in \Sigma^+} |\alpha(H)|\, d\bar{k} dH.$$

On the other hand we have the following lemma.

LEMMA 37. *Let $\mathfrak{b}_0$ be a connected component of $\mathfrak{a}'_{\mathfrak{p}_0}$. Then*

$$\mathfrak{a}'_{\mathfrak{p}_0} = \bigcup_{s \in W} s\mathfrak{b}_0.$$

Let $\bar{\mathfrak{b}}_0$ denote the closure of $\mathfrak{b}_0$ in $\mathfrak{a}_{\mathfrak{p}_0}$. Since $\mathfrak{b}_0$ is obviously closed

in $\mathfrak{a}'_{\mathfrak{p}_0}$, $\bar{\mathfrak{b}}_0 \cap \mathfrak{a}'_{\mathfrak{p}_0} = \mathfrak{b}_0$. Consider $\phi(\bar{K}_0 \times \mathfrak{b}_0) \subset \mathfrak{p}'_0$. Since ϕ is regular on $\bar{K}_0 \times \mathfrak{a}'_{\mathfrak{p}_0}$, $\phi(\bar{K}_0 \times \mathfrak{b}_0)$ is open. Moreover since $\bar{K}_0$ is compact it follows easily that $\phi(\bar{K}_0 \times \bar{\mathfrak{b}}_0)$ is closed in $\mathfrak{p}_0$. Hence $\phi(\bar{K}_0 \times \mathfrak{b}_0) = \phi(\bar{K}_0 \times \bar{\mathfrak{b}}_0) \cap \mathfrak{p}'_0$ is closed in $\mathfrak{p}'_0$. But we know that $\mathfrak{p}'_0$ is connected and therefore $\phi(\bar{K}_0 \times \mathfrak{b}_0) = \mathfrak{p}'_0$. Hence in particular if $H \in \mathfrak{a}'_{\mathfrak{p}_0}$, $H = \mathrm{Ad}(k)H_0$ for some $k \in K_0$ and $H_0 \in \mathfrak{b}_0$. But then it follows from Lemma 35 that $k \in M'$ and therefore $H = sH_0$ for some $s \in W$. This proves that $\mathfrak{a}'_{\mathfrak{p}_0} \subset \bigcup_{s \in W} s\mathfrak{b}_0$. Since the reverse inclusion is obvious we get the lemma.

COROLLARY. *$\mathfrak{a}'_{\mathfrak{p}_0}$ has only a finite number of connected components.*

Now as we have seen earlier, $\prod_{\alpha \in \Sigma^+} |\alpha(H)|^2$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$ is the determinant of the linear transformation in $\mathfrak{p}/\mathfrak{a}_{\mathfrak{p}}$ corresponding to $(adH)^2$. From this it follows that the value of this expression does not change if we replace H by sH $(s \in W)$. Therefore it is obvious from Lemma 37 that

$$\int_{K_0 \times \mathfrak{B}_{\mathfrak{p}_0}} f(\mathrm{Ad}(k)H) \prod_{\alpha \in \Sigma^+} |\alpha(H)| \, dk dH = w \int_{K_0 \times \mathfrak{b}_0} f(\mathrm{Ad}(k)H) \prod_{\alpha \in \Sigma^+} |\alpha(H)| \, dk dH$$

where w is the number of connected components of $\mathfrak{a}'_{\mathfrak{p}_0}$, $\mathfrak{b}_0$ is any such component and f is a continuous function on $\mathfrak{p}_0$ vanishing outside a compact set. Combining this with Lemma 31 and its corollary, we obtain the following result.

LEMMA 38. *It is possible to normalize the Haar measure dx on G_0 in such a way that the following condition is fulfilled. If f is a continuous function on G_0 vanishing outside a compact set and $\mathfrak{b}_0$ is any connected component of $\mathfrak{a}'_{\mathfrak{p}_0}$,*

$$\int_{G_0} f(x)\,dx = \int_{K_0 \times \mathfrak{p}_0} f(k \exp X) \det(\sinh \mathrm{ad}\, X / \mathrm{ad}\, X)_{\mathfrak{p}}\, dk dX$$
$$= cw \int_{\mathfrak{b}_0} \prod_{\alpha \in \Sigma^+} |e^{\alpha(H)} - e^{-\alpha(H)}|\, dH \int_{K_0 \times K_0} f(k \exp H\, k')\, dk dk'$$

where w is the number of connected components of $\mathfrak{a}'_{\mathfrak{p}_0}$ and c is a positive constant given by the relation

$$\int_{\mathfrak{p}_0} e^{-\|X\|^2} \det(\sinh adX/adX)_{\mathfrak{p}}\, dX = c \int_{\mathfrak{B}_{\mathfrak{p}_0}} e^{-\|H\|^2} \prod_{\alpha \in \Sigma^+} |e^{\alpha(H)} - e^{-\alpha(H)}|\, dH.$$

For the proof we have only to notice the fact that

$$\det(\sinh adH/adH) = \prod_{\alpha \in \Sigma^+} (\sinh \alpha(H)/\alpha(H)) \qquad (H \in \mathfrak{a}_{\mathfrak{p}_0}).$$

The following lemma is also due to Cartan [2(a)].

LEMMA 39. *Let* $'\mathfrak{a}_{\mathfrak{p}_0}$ *be any maximal abelian subspace of* $\mathfrak{p}_0$. *Then*

$$'\mathfrak{a}_{\mathfrak{p}_0} = \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$$

for some $k \in K_0$.

Define $\mathfrak{a}'_{\mathfrak{p}_0}$ as above and choose $H_0 \in \mathfrak{a}'_{\mathfrak{p}_0}$. Then $\mathfrak{a}_{\mathfrak{p}_0}$ is exactly the centralizer of H_0 in $\mathfrak{p}_0$. If we apply Lemma 33 to $'\mathfrak{a}_{\mathfrak{p}_0}$ (instead of $\mathfrak{a}_{\mathfrak{p}_0}$), it follows that $H_0 \in \mathrm{Ad}(k^{-1})'\mathfrak{a}_{\mathfrak{p}_0}$ for some $k \in K_0$. Since $'\mathfrak{a}_{\mathfrak{p}_0}$ is abelian the same holds for $\mathrm{Ad}(k^{-1})'\mathfrak{a}_{\mathfrak{p}_0}$ and therefore $\mathrm{Ad}(k^{-1})'\mathfrak{a}_{\mathfrak{p}_0} \subset \mathfrak{a}_{\mathfrak{p}_0}$ or $'\mathfrak{a}_{\mathfrak{p}_0} \subset \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$. But then since $'\mathfrak{a}_{\mathfrak{p}_0}$ is maximal abelian in $\mathfrak{p}_0$ it is obvious that $'\mathfrak{a}_{\mathfrak{p}_0} = \mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0}$.

Let $\mathfrak{A}$ be the analytic subgroup of G_0 corresponding to $\mathfrak{a}_{\mathfrak{p}_0}$. We conclude from Lemma 31 that $\mathfrak{A}$ is closed and $H \to \exp H$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$ is a topological mapping of $\mathfrak{a}_{\mathfrak{p}_0}$ onto $\mathfrak{A}$. Moreover as we have seen in Section 8,

$$\begin{aligned} |D(\exp H)| &= \prod_{\alpha \in \Sigma} |e^{\alpha(H)} - e^{-\alpha(H)}| \\ &= \prod_{\alpha \in \Sigma^+} |e^{\alpha(H)} - e^{-\alpha(H)}|^2 \qquad (H \in \mathfrak{a}_{\mathfrak{p}_0}). \end{aligned}$$

Therefore if we define $\mathfrak{A}'$ as in Lemma 22, it is clear that $\mathfrak{A}' = \exp(\mathfrak{a}'_{\mathfrak{p}_0})$ and the number of connected component of $\mathfrak{A}'$ and $\mathfrak{a}'_{\mathfrak{p}_0}$ is the same. Moreover it follows from Lemma 39 without difficulty that the numbers c and w of Lemma 38 are independent of the choice of $\mathfrak{a}_{\mathfrak{p}_0}$. Therefore the second statement of Lemma 22 is now obvious.

It remains to prove the first part of Lemma 22. Let

$$e(X) = \exp(-1)^{\frac{1}{2}}X \in U \qquad (X \in \mathfrak{p}_0).$$

LEMMA 40. $\mathfrak{A}^* = e(\mathfrak{a}_{\mathfrak{p}_0})$ *is compact.*

Since U is compact it is enough to prove that $\mathfrak{A}^*$ is closed in U. Let $\bar{\mathfrak{A}}^*$ denote the closure of $\mathfrak{A}^*$ in U. Obviously $\mathfrak{A}^*$ and therefore $\bar{\mathfrak{A}}^*$ is a connected Lie subgroup of U. G_c being simply connected, we "extend" θ to a (complex) automorphism of G_c so that $\theta(\exp X) = \exp\theta(X)$ $(X \in \mathfrak{g})$. Since $\theta(H) = -H$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$, it is clear that $\theta(a) = a^{-1}$ for $a \in \mathfrak{A}^*$ and therefore by continuity also for $a \in \bar{\mathfrak{A}}^*$. Hence if X is an element in $\mathfrak{u}$ which lies in the Lie algebra of $\bar{\mathfrak{A}}^*$, $\theta(X) = -X$ and therefore $X \in \mathfrak{u} \cap \mathfrak{p} = (-1)^{\frac{1}{2}}\mathfrak{p}_0$. But since $\mathfrak{a}_{\mathfrak{p}_0}$ is maximal abelian in $\mathfrak{p}_0$, it is obvious that $X \in (-1)^{\frac{1}{2}}\mathfrak{a}_{\mathfrak{p}_0}$. This proves that $\mathfrak{A}^*$ and $\bar{\mathfrak{A}}^*$ have the same Lie algebra and therefore $\mathfrak{A}^* = \bar{\mathfrak{A}}^*$.

Define $\mathfrak{A}^{*\prime}$ as in Lemma 22. Then $e(H) \in \mathfrak{A}^{*\prime}$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$ if and only if

$$\prod_{\alpha \in \Sigma^+} |\sin \alpha(H)| \neq 0.$$

Put $U_* = U/K_0$ and $u_* = uK_0$ $(u \in U)$. We define an analytic mapping ψ of $\bar{K}_0 \times \mathfrak{A}^*$ into U_* as follows:

$$\psi(\bar{k}, a) = (ka)_* \qquad (\bar{k} \in \bar{K}, a \in \mathfrak{A}^*)$$

where k is any element in the coset $\bar{k}$.

LEMMA 41. *ψ is regular and open on $\bar{K}_0 \times \mathfrak{A}^{*\prime}$ and $\psi(\bar{K} \times \mathfrak{A}^*) = U_*$.*

Let u be an element in U and f a function which is defined and analytic on some neighborhood of u in U. Suppose $u = ka$ where $k \in K_0$ and $a \in \mathfrak{A}^*$. Then if $H \in (-1)^{\frac{1}{2}}\mathfrak{a}_{\mathfrak{p}_0}$ and $X \in k_0$,

$$\left\{ \frac{d}{dt} f(ka \exp tH) \right\}_{t=0} = (Hf)(u)$$

$$\left\{ \frac{d}{dt} f(k \exp tX\, a) \right\}_{t=0} = (X'f)(u)$$

where $X' = \mathrm{Ad}(a^{-1})X$. Let T denote the linear mapping of $\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}}$ into $\mathfrak{g}$ given by

$$H \to H \qquad (H \in \mathfrak{a}_{\mathfrak{p}})$$

$$X \to \mathrm{Ad}(a^{-1})X \qquad (X \in \mathfrak{k}).$$

Since $\mathfrak{m}$ lies in the kernel of T, we get a linear mapping $\bar{T}$ of $(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m}$ into $\mathfrak{g}/\mathfrak{k}$ by going over to the factor spaces. Now suppose $T(H + X) \in \mathfrak{k}$ $(H \in \mathfrak{a}_{\mathfrak{p}}, X \in \mathfrak{k})$. Then $H + \mathrm{Ad}(a^{-1})X \in \mathfrak{k}$. But

$$\mathrm{Ad}(a^{-1})X \equiv \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))X \bmod \mathfrak{k}$$

and $(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))X \in \mathfrak{p}$. Hence

$$H + \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))X \in \mathfrak{k} \cap \mathfrak{p} = 0.$$

However it is easy to check that $(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))\mathfrak{g}$ and $\mathfrak{a}_{\mathfrak{p}}$ are orthogonal. Therefore $H = 0$ and $\frac{1}{2}(Ad(a^{-1}) - Ad(a))X = 0$. Now if in particular $a \in \mathfrak{A}^{*\prime}$, this clearly implies that $X \in \mathfrak{m}$. This shows that if $a \in \mathfrak{A}^{*\prime}$, the kernel of $\bar{T}$ is zero. Moreover as we have seen before $\mathfrak{g} = \mathfrak{k} + \mathfrak{a}_{\mathfrak{p}} + \mathfrak{n}^+ = \mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+$ and therefore

$$(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m} \cong \mathfrak{g}/(\mathfrak{m} + \mathfrak{n}^+) \cong \mathfrak{p} \cong \mathfrak{g}/\mathfrak{k}$$

Hence $(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m}$ and $\mathfrak{g}/\mathfrak{k}$ have the same dimension. This proves that $\bar{T}$ is an isomorphism of $(\mathfrak{k} + \mathfrak{a}_{\mathfrak{p}})/\mathfrak{m}$ onto $\mathfrak{g}/\mathfrak{k}$ and therefore ψ is open and regular on $\bar{K}_0 \times \mathfrak{A}^{*\prime}$. Let du_* and $d\bar{k}$ denote the invariant measures on the

homogeneous spaces U_* and $\bar{K}_0$ respectively. We assume that

$$\int_{U_*} du_* = \int_{\bar{K}_0} d\bar{k} = 1.$$

Since $\mathfrak{g} = \mathfrak{n}^- + \mathfrak{a}_\mathfrak{p} + \mathfrak{m} + \mathfrak{n}^+$,

$$\mathfrak{g}/(\mathfrak{m} + \mathfrak{n}^+) \cong \mathfrak{a}_\mathfrak{p} + \mathfrak{n}^-$$

and therefore if we identify $(\mathfrak{k} + \mathfrak{a}_\mathfrak{p})/\mathfrak{m}$ and $\mathfrak{g}/\mathfrak{k}$ with $\mathfrak{a}_\mathfrak{p} + \mathfrak{n}^-$ under the isomorphism indicated above, $\bar{T}$ may be regarded as a linear mapping of $\mathfrak{a}_\mathfrak{p} + \mathfrak{n}^-$ into itself. If $Z \in \mathfrak{n}^-$,

$$\begin{aligned} \mathrm{Ad}(a^{-1})(Z + \theta(Z)) &\equiv \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))(Z + \theta(Z)) \\ &\equiv Z' - \theta(Z') \bmod \mathfrak{k} \end{aligned}$$

where $Z' = \frac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))Z$. From this it follows easily that

$$\bar{T}(H + Z) = H + \tfrac{1}{2}(\mathrm{Ad}(a^{-1}) - \mathrm{Ad}(a))Z \qquad (H \in \mathfrak{a}_\mathfrak{p}, Z \in \mathfrak{n}^-)$$

and therefore

$$|\det \bar{T}| = 2^{-q} |D(a)|^{\frac{1}{2}}$$

where q is the number of roots in Σ^+. This proves that

$$du_* \sim |D(a)|^{\frac{1}{2}} da d\bar{k} \qquad (u_* = \psi(\bar{k}, a))$$

if $\bar{k} \in \bar{K}_0$ and $a \in \mathfrak{A}^{*\prime}$ and da denotes the Haar measure on $\mathfrak{A}^*$. We shall need this result presently.

For each $\alpha \in \Sigma^+$ we consider the character ξ_α of $\mathfrak{A}^*$ given by $\xi_\alpha(e(H)) = \exp((-1)^{\frac{1}{2}}\alpha(H))$ $(H \in \mathfrak{a}_{\mathfrak{p}_0})$. Then obviously $a \to \xi_\alpha(a^2)$ $(a \in \mathfrak{A}^*)$ is also a non-trivial character of $\mathfrak{A}^*$. Hence its kernel $\mathfrak{A}^*_\alpha$ is a closed subgroup of $\mathfrak{A}^*$ and $\dim \mathfrak{A}^*_\alpha = \dim \mathfrak{A}^* - 1$. Let M_α be the subgroup of K_0 consisting of all $k \in K_0$ such that $(ka)_* = a_*$ for every $a \in \mathfrak{A}^*_\alpha$. Obviously M_α is closed and it contains M. Now consider the elements Y_α, Z_α of Lemma 34. If $H \in \mathfrak{a}_{\mathfrak{p}_0}$ it is obvious that

$$\mathrm{Ad}(e(H))Z_\alpha = \cos \alpha(H) Z_\alpha + (-1)^{\frac{1}{2}} \sin \alpha(H) Y_\alpha.$$

Moreover if $e(H) \in \mathfrak{A}^*_\alpha$, $\alpha(H)/\pi$ is an integer and therefore

$$\mathrm{Ad}(a) Z_\alpha = \pm Z_\alpha$$

for $a \in \mathfrak{A}^*_\alpha$. From this we conclude immediately that Z_α lies in the Lie algebra of M_α. Since $Z_\alpha \notin \mathfrak{m}_0$, $\dim M_\alpha > \dim M$. But $\psi(\bar{K}_0 \times \mathfrak{A}^*_\alpha)$ may be regarded as the image of $(K_0/M_\alpha) \times \mathfrak{A}^*_\alpha$ under an analytic mapping and therefore it follows from Lemma 32 that

$$\dim \psi(\bar{K}_0 \times \mathfrak{A}^*_a) \leqq \dim K_0/M_a + \dim \mathfrak{A}^*_a$$
$$\leqq \dim \bar{K}_0 + \dim \mathfrak{A}^* - 2 = \dim U_* - 2.$$

Hence if $V = \bigcup_{a \in \Sigma^+} \psi(\bar{K}_0 \times \mathfrak{A}^*_a)$, V is compact and $\partial \dim V \leqq \dim U_* - 2$ and therefore (see Hurewicz and Wallman [6, p. 48]) the complement ${}^cV$ of V in U_* is connected. Moreover ${}^cV$ is a dense subset of U_*. Now consider

$$U'_* = {}^cV \cap \psi(\bar{K}_0 \times \mathfrak{A}^*).$$

Since $\bar{K}_0 \times \mathfrak{A}^*$ is compact, U'_* is closed in ${}^cV$. On the other hand V and $\psi(\bar{K}_0 \times \mathfrak{A}^{*\prime})$ cannot have any point in common (see Corollary 1 to Lemma 42 below) and therefore

$$U'_* = \psi(\bar{K}_0/\mathfrak{A}^{*\prime}).$$

Since ψ is open on $\bar{K}_0 \times \mathfrak{A}^{*\prime}$, U'_* is a nonempty open subset of U_*. Therefore ${}^cV$ being connected, we can conclude that $U'_* = {}^cV$. This shows that the compact set $\psi(\bar{K}_0 \times \mathfrak{A}^*)$ is dense in U_*. Hence $\psi(\bar{K}_0 \times \mathfrak{A}^*) = U_*$ and Lemma 41 is proved.

For any $a \in \mathfrak{A}^*$ and $s \in W$, define $a^s = kak^{-1}$ where k is some element of M' lying in the coset s. Then $a^s \in \mathfrak{A}^*$ and $|D(a^s)| = |D(a)|$. Hence $(\mathfrak{A}^{*\prime})^s = \mathfrak{A}^{*\prime}$.

LEMMA 42. *Suppose* $(ka_1)_* = (a_2)_*$ $(k \in K_0; a_1, a_2 \in \mathfrak{A}^*)$. *Then if* $a_1 \in \mathfrak{A}^{*\prime}$, k *lies in* M' *and* a_2 *in* $\mathfrak{A}^{*\prime}$.

For $ka_1 = a_2k'$ where $k' \in K_0$. Applying the automorphism θ to both sides we get $ka_1^{-1} = a_2^{-1}k'$ and therefore $ka_1^2k^{-1} = a_2^2$. Since $a_1 \in \mathfrak{A}^{*\prime}$, $\mathfrak{a}_{\mathfrak{p}_0}$ is exactly the set of all elements $H \in \mathfrak{p}_0$ such that $\mathrm{Ad}(a_1^2)H = H$. Therefore it follows from the above relation that $\mathrm{Ad}(k)\mathfrak{a}_{\mathfrak{p}_0} = \mathfrak{a}_{\mathfrak{p}_0}$ and so $k \in M'$. Also it is now clear that $a_2 \in \mathfrak{A}^{*\prime}$.

COROLLARY 1. *Suppose* $a \in \mathfrak{A}^{*\prime}$ *and* $\bar{k} \in \bar{K}_0$. *Then if* w_0 *is the order of* W, $\psi(\bar{k}, a)$ *has exactly* w_0 *distinct preimages in* $\bar{K}_0 \times \mathfrak{A}^*$ *namely* $(\bar{k}s^{-1}, a^s)$ $(s \in W)$.

Choose an element k in the coset $\bar{k}$. Then if $\psi(\bar{k}_1, a_1) = \psi(\bar{k}, a)$ $(k_1 \in K_0$, $a_1 \in \mathfrak{A}^*)$, it follows from Lemma 42 that $k^{-1}k_1 \in M'$ and therefore if s^{-1} denotes the corresponding element of W, $\bar{k}_1 = \bar{k}s^{-1}$ and $a_1 = a^s$.

COROLLARY 2. *There exists a positive number* c *with the following property. If* f *is any continuous function on* U_*,

$$\int_{U_*} f(u_*)\,du_* = c \int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da \int_{\bar{K}_0} f(\psi(\bar{k}, a))\,d\bar{k}.$$

We have seen that $du_* \sim |D(a)|^2 \, da d\bar{k}$ if $u_* = \psi(\bar{k}, a)$ and $\bar{k} \in \bar{K}_0$, $a \in \mathfrak{A}^{*\prime}$. Now if we define $V = \bigcup_{\alpha \in \Sigma^+} \psi(\bar{K}_0 \times \mathfrak{A}^*_\alpha)$ as before, it follows from Lemma 45 (Section 13) that V is a null set on U_* with respect to the measure du_*. Similarly $\bigcup_{\alpha \in \Sigma^+} \mathfrak{A}^*_\alpha$ is a null set on $\mathfrak{A}^*$ with respect to the measure da. Since $\psi(\bar{K}_0 \times \mathfrak{A}^{*\prime}) = {}^cV$, our assertion follows from Corollary 1 above.

Put $J^* = K_0 \cap \mathfrak{A}^*$. Since $\mathfrak{k}_0 \cap (-1)^{\frac{1}{2}} \mathfrak{a}_{\mathfrak{p}_0} = 0$, J^* is both discrete and compact and so it is finite. It is obvious that $(J^*)^s = J^*$ $(s \in W)$.

LEMMA 43. *Let B^*_0 be any connected component of $\mathfrak{A}^{*\prime}$. Then*

$$\mathfrak{A}^{*\prime} = \bigcup_{s \in W} (B^*_0)^s J^*.$$

Hence $\mathfrak{A}^{*\prime}$ has only a finite number of connected components.

Let $\bar{B}^*_0$ denote the closure of B^*_0 in $\mathfrak{A}^*$. Since B^*_0 is closed in $\mathfrak{A}^{*\prime}$, $B^*_0 = \bar{B}^*_0 \cap \mathfrak{A}^{*\prime}$. Now define V and ${}^cV$ as above. Then

$$\psi(\bar{K}_0 \times B^*_0) = {}^cV \cap \psi(\bar{K}_0 \times \bar{B}^*_0)$$

is both open and closed in ${}^cV$ and therefore, ${}^cV$ being connected,

$${}^cV = \psi(\bar{K}_0 \times B^*_0) \text{ and } U_* = \psi(\bar{K}_0 \times \bar{B}^*_0).$$

Now suppose $a \in \mathfrak{A}^{*\prime}$. Then we can find elements $b \in \bar{B}^*_0$ and $k \in K_0$ such that $a_* = (kb)_*$. Hence $(k^{-1}a)_* = b_*$ and therefore $k \in M'$ from Lemma 42. This means that $b_* = (a^s)_*$ for some $s \in W$ and $b^{-1}a^s \in K_0 \cap \mathfrak{A}^* = J^*$. Conversely if $z \in J^*$, $z^{-1} = \theta(z)$ and so $z^2 = 1$. Hence $|D(az)| = |D(a)|$ for any $a \in \mathfrak{A}^*$. This shows that $\bigcup_{s \in W} (B^*_0)^s J^*$ is contained in $\mathfrak{A}^{*\prime}$ and so the lemma follows.

COROLLARY. *Let w^* be the number of connected components of $\mathfrak{A}^{*\prime}$. Then if f is any continuous function on U,*

$$\int_U f(u)\,du \int_{\mathfrak{A}^*} |D(a)|^2 \, da = w^* \int_{B_0^*} |D(a)|^2 \, da \int_{K_0 \times K_0} f(kak')\,dk dk'.$$

Let f^* denote the function on U_* given by

$$f^*(u_*) = \int_{K_0} f(uk)\,dk \qquad (u \in U).$$

Then

$$\int_{U_*} f^*(u_*)\,du_* = \int_U f(u)\,du$$

and

$$\int_{K_0} f(kak')\,dk' = f^*((ka)_*).$$

Therefore

$$\int_{K_0\times K_0} f(kak')\,dkdk' = \int_{K_0} f^*((ka_*)\,dk$$

and so it follows from Corollary 2 to Lemma 42 that

$$\int_U f(u)\,du = c\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kak')\,dkdk'.$$

Now we can replace the integral on $\mathfrak{A}^*$ on the right by the corresponding integral on $\mathfrak{A}^{*\prime}$. We know that any component B^* of $\mathfrak{A}^{*\prime}$ is of the form $(B^*_0)^s z$ $(s\in W, z\in J^*)$. Therefore if m is any element of M' in the coset s, it is clear that

$$\begin{aligned}\int_{B^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kak')\,dkdk' &= \int_{B_0^*} |D(a^s z)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kmam^{-1}zk')\,dkdk' \\ &= \int_{B_0^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kak')\,dkdk'\end{aligned}$$

since $|D(a^s z)| = |D(a)|$. This shows that

$$\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kak')\,dkdk' = w^* \int_{B_0^*} |D(a)|^{\frac{1}{2}}\,da \int_{K_0\times K_0} f(kak')\,dkdk'.$$

In particular if we take the function $f=1$, we get

$$1 = \int_U du = c\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da$$

and this gives our result.

Now if we normalize the Haar measure da on $\mathfrak{A}^*$ in accordance with Lemma 22, it follows from Lemma 39 that the numbers w^* and $\int_{\mathfrak{A}^*} |D(a)|^{\frac{1}{2}}\,da$ are independent of the choice of $\mathfrak{a}_{p_0}$. This completes the proof of Lemma 22.

13. Appendix. We shall now give a proof of Lemma 32. Let m and n be the dimensions of M and N respectively and let $(t_1,\cdots,t_m)$ be a coordinate system on M which is valid on some open neighborhood V of a point $x_0\in M$. We assume $t_i(x_0)=0$ $i=1,\cdots,m$. For any positive integer p let R^p denote the Cartestian product of R with itself p times. It is obvious that we can chose a compact neighborhood W of x_0 and a positive number a such that $W\subset V$ and the mapping $x\to(t_1(x),\cdots,t_m(x))$ $(x\in W)$ maps W topologically on the cube $|t_i|\leqq a$ in R^m. Any set W defined in this way

is called a cubic set in M (with respect to the coordinate system $(t_1, \cdots, t_m)$). Cubic sets in N are defined similarly.

Since M satisfies the countability axioms (see Chevalley [3]) we can find a countable family V_i $(i = 1, 2, \cdots)$ of cubic sets in M such that the following conditions hold: (1) $M = \bigcup_i V_i$ and (2) for each i, $f(V_i)$ is contained in some cubic set in N. Therefore, in view of the sum theorem of dimension theory (Hurewicz and Wallman [6, p. 30]) it is enough to show that $\partial\dim f(V_i) \leqq m$ for each i. This however is an immediate consequence of the lemma below.

We regard R^p as an additive group in the usual way and if

$$a = (a_1, \cdots, a_p) \in R^p, \text{ we put } |a| = \Big(\sum_{1 \leqq i \leqq p} a_i^2\Big)^{\frac{1}{2}}.$$

Let I be the unit interval $0 \leqq t \leqq 1$ in R and let I^m be the Cartesian product of I with itself m times. Then $I^m \subset R^m$.

LEMMA 44. *Let f be a mapping of I^m into R^n such that*

$$|f(b) - f(a)| \leqq c\,|b - a| \qquad (a, b \in I^m)$$

where c is a fixed number. Then $\partial\dim f(I^m) \leqq m$.

It is enough to prove that the Hausdorff $(m+1)$-measure of $f(I^m)$ is zero (see Hurewicz and Wallman [6, p. 104]). But this follows at once from our hypothesis on f and the fact that the Hausdorff $(m+1)$-measure of I^m is zero (Hurewicz and Wallman [6, p. 103]).

Let μ be a Borel measure on N. We say that μ is locally euclidean if the following condition holds. For each x_0 in N we can find a cubic set W in N with respect to a coordinate system $(t_1, \cdots, t_n)$ such that (1) x_0 lies in the interior of W and (2) μ is completely continuous (on W) with respect to the Euclidean measure $dt = dt_1 \cdots dt_n$ on W.

LEMMA 45. *Let N be a manifold satisfying the countability axioms and let μ be a locally euclidean measure on N. Then if Q is any subset of N with $\partial\dim Q < \partial\dim N$, Q is a null set with respect to μ.*

We can cover N by a countable family of cubic sets V_i $i = 1, 2, \cdots$. For any fixed i, choose a coordinate system $(t_1, \cdots, t_n)$ valid on some open neighborhood of V_i such that V_i is a cubic set with respect to this system. Obviously it would be enough to prove that $Q \cap V_i$ is a null set with respect to the measure $dt = dt_1 \cdots dt_n$. So we have to prove the following lemma.

LEMMA 46. *Let Q be a subset of I^n such that $\partial \dim Q < n$. Then Q is a null set with respect to the Euclidean measure on I^n.*

Since $\partial \dim Q < n$, the Hausdorff n-measure of Q is zero. Our assertion therefore follows immediately from the definition of the Hausdorff measure.

THE INSTITUTE FOR ADVANCED STUDY,
PRINCETON, N. J.

COLUMBIA UNIVERSITY,
NEW YORK, N. Y.

REFERENCES.

[1] V. Bargmann, *Ann. of Math.*, vol. 48 (1947), pp. 568-640.

[2] E. Cartan, (a) *Ann. École Norm. Sup.*, vol. 44 (1927), pp. 345-467.
(b) *J. Math. Pures Appl.*, vol. 8 (1929), pp. 1-33.
(c) *Abh. Math. Seminar Hamburg*, vol. 11 (1925), pp. 116-162.

[3] C. Chevalley, *Theory of Lie groups*, Princeton University Press, 1946.

[4] R. Godement, *C. R. Acad. Sci. Paris*, vol. 225 (1947), (a) pp. 521-523, (b) pp. 657-659.

[5] Harish-Chandra. (a) *Trans. Amer. Math. Soc.*, vol. 70 (1951), pp. 28-96.
(b) vol. 75 (1953), pp. 185-243.
(c) vol. 76 (1953), pp. 234-253.
(d) *Proc. Nat. Acad. Sci.*, vol. 41 (1955), pp. 314-317.
(e) *Amer. Jour. Math.*, vol. 77 (1955), pp. 743-777.
(f) *Amer. Jour. Math.*, vol. 78 (1956), pp. 1-41.

[6] W. Hurewicz and H. Wallman, *Dimension theory*, Princeton University Press, 1948.

[7] J. L. Koszul, *Bull. Soc. Math.*, vol. 78 (1950), pp. 65-127.

[8] F. I. Mautner, *Ann. of Math.*, vol. 52 (1950), pp. 528-556.

[9] G. D. Mostow, *Bull. Amer. Math. Soc.*, vol. 55 (1949), pp. 969-980.

[10] I. E. Segal, (a) *Ann. of Math.*, vol. 52 (1950), pp. 272-292.
(b) *Proc. Amer. Math. Soc.*, vol. 3 (1952), pp. 13-15.

[11] H. Weyl, (a) *Math. Zeit.*, vol. 24 (1925), pp. 328-395.
(b) The structure and representations of continuous groups, Princeton, The Institute for Advanced Study, 1935.

LIE AND JORDAN SYSTEMS IN SIMPLE RINGS WITH INVOLUTION.*

By I. N. Herstein.

In previous papers, ([3], [4], [5]), the author has recently considered the question of Lie and Jordan simplicity of simple, associative rings and of certain subsets thereof. Baxter [1] completed certain cases, namely characteristic 2 and 3. These considerations were motivated by the classical results (in the theory of simple Lie algebras) which yielded that total matrix algebras over fields of charateristic 0 gave rise, in a natural way, to simple Lie algebras.

However, in the total matrix algebras over fields, adjoints (involutions) can be defined in a variety of ways, and it was also known classically that the skew elements under the adjoint also led to simple Lie algebras; more recently, it has been demonstrated that the self-adjoint elements yield simple Jordan algebras. With these examples as motivation, we proceed, in this paper, to study the general situation, namely the Lie and Jordan structure of the skew and self-adjoint elements of arbitrary simple rings possessing involutions.

In the course of this study several other results, interesting in their own rights, fall our way. For instance we prove (Theorem 9) that the subring generated by the self-adjoint elements is A except in the 4-dimensional case; this generalizes a result of Dieudonné [2]. We also show that under the same conditions, as above, the skew elements generate the full ring (Theorem 15); this is considerably more difficult than the case of the self-adjoint elements. There are also some special results about representations in various forms of self-adjoint elements and also concerning the taking of commutators twice of the skew elements.

1. Preliminaries. Let A be a simple ring of characteristic different from 2. We say a mapping, *, from A to A is an adjoint or involution on A if

* Received November 14, 1955.

(1) $a^{**} = a$, (2) $(a+b)^* = a^* + b^*$, (3) $(ab)^* = b^*a^*$,

for all a, b in A.

As usual, we define S, the set of self-adjoint elements of A, by $S = \{x \in A \mid x^* = x\}$. Clearly S is an additive subgroup of A. Moreover, if $a, b \in S$ then $ab + ba \in S$. Thus S is a Jordan subring of A under the Jordan product which is defined in A by $x \circ y = xy + yx$ for $x, y \in A$. An additive subgroup, U, of S is said to be a *Jordan ideal* of S if whenever $u \in U$, $x \in S$ then $ux + xu \in U$.

We further define K, the set of skew elements of A, by $K = \{x \in A \mid x^* = -x\}$. As is readily verified, K is an additive subgroup of A, and if $x, y \in K$ then $xy - yx \in K$. Thus K is a Lie subring of A under the Lie product $[x, y]$ defined in A by $[x, y] = xy - yx$ for all $x, y \in A$. We say that an additive subgroup, U, of K is a *Lie ideal* of K if $x \in K$, $u \in U$ implies that $xu - ux \in U$.

Certain facts about adjoints are trivial and we shall make use of these throughout the body of this paper without proof, explanation or reference; such facts include, for instance, $x^* - x \in K$, $x^* + x \in S$, $s \in S$, $k \in K$ then $sk + ks \in K$, $sk - ks \in S$, etc.

2. Jordan simplicity of S. The purpose of this section is to show that the results from matrix theory extend to general simple rings, that is, that S is a simple Jordan ring. Throughout this section U will denote a non-zero Jordan ideal of S. Our purpose is to show that $U = S$.

LEMMA 1. *If $x, y \in S$ and $u \in U$ then $(xu - ux)y - y(xu - ux) \in U$.*

Proof. Since U is a Jordan ideal of S and since $xy + yx \in S$, it follows that $(xy + yx)u + u(xy + yx) \in U$. However,

$$(xy + yx)u + u(xy + yx)$$
$$= \{x(yu + uy) + (yu + uy)x\} + \{(ux + xu)y - y(ux - xu)\}.$$

Now $yu + uy \in U$ since U is a Jordan ideal of S; thus, by the same token, $x(uy + yu) + (uy + yu)x \in U$. That is, the first $\{\ \}$ on the right-hand side of the relation above is in U. Since the left-hand side of this identity is also in U, we are left with the fact that the second $\{\ \}$ on the right-hand side must be in U. That is, $(ux - xu)y - y(ux - xu)$ is in U, which is the required lemma.

LEMMA 2. *If $u \in U$, $b \in K$ then $bu^2 - u^2b \in U$.*

For, $ub - bu \in S$, and so $(ub - bu)u + u(ub - bu) \in U$; since

$$(ub - bu)u + u(ub - bu) = u^2b - bu^2$$

this completes the proof of the lemma.

LEMMA 3. *If $u \in U$, $x \in S$ then $xux \in U$.*

Proof. $2xux = \{x(xu + ux) + (xu + ux)x\} - \{x^2u + ux^2\}$. Since U is a Jordan ideal of S, and since $x^2 \in S$ along with x, both $\{\ \}$ on the right-hand side are in U, and so $2xux \in U$. Thus $2(2xux) \in U$; that is $(2x)u(2x) \in U$ for all $x \in S$. Since the characteristic of A is not 2, it is easily seen that $2S = S$; hence $(2x)u(2x) \in U$ for all $x \in S$ implies $xux \in U$ for all $x \in S$.

COROLLARY. *If $x, y \in S$ and $u \in U$ then $xuy + yux \in U$.*

For, linearizing the result of the lemma by replacing x by $x + y$ the corollary follows immediately from the lemma.

Our aim is to show that $au^4b + b^*u^4a^*$ is in U for all $a, b \in A$. We have, in the corollary above, disposed of the case $a^* = a$, $b^* = b$. We proceed in the next few lemmas to consider the other special possibilities, $a^* = -a$, $b^* = +b$, $a^* = -a$, $b^* = -b$, and then to combine them for the result at the opening of this paragraph. We now show

LEMMA 4. *If $u \in U$, $b \in S$, $a \in K$ then $bu^2a - au^2b \in U$.*

Proof. By Lemma 2, since $a \in K$, $u^2a - au^2 \in U$. Then, U being a Jordan ideal of S, $b(u^2a - au^2) + (u^2a - au^2)b \in U$. Now

$$(1) \qquad b(u^2a - au^2) + (u^2a - au^2)b = bu^2a - au^2b - bau^2 + u^2ab.$$

Consider

$$(2) \qquad a(u^2b - bu^2) + (u^2b - bu^2)a = au^2b - bu^2a - abu^2 + u^2ba.$$

Adding (1) and (2), the right-hand sides add up to $u^2(ab + ba) - (ab + ba)u^2$; since $ab + ba \in K$, by Lemma 2, $u^2(ab + ba) - (ab + ba)u^2$ is in U. Thus the sum of the right-hand sides of (1) and (2) is in U; consequently the sum of the left-hand sides is also in U. Since the left-hand side of (1) is already in U, we obtain that the left-hand side of (2) must also be in U. Thus, if we now subtract (1) from (2) we observe that the elements we get on each side of the resulting equation are in U. In particular, the difference of the right-hand sides is an element of U; whence

$$(3) \qquad 2(au^2b - bu^2a) + (ba - ab)u^2 + u^2(ba - ab) \in U.$$

Now $2u^2 = uu + uu \in U$, so $2u^2s + s2u^2 = u^2(2s) + (2s)u^2 \in U$ for all $s \in S$, and since $2S = S$, we have $u^2s + su^2 \in U$. In particular, in (3), since $ab - ba \in S$, $u^2(ab - ba) + (ab - ba)u^2 \in U$; from which (3) reduces to $2(au^2b - bu^2a) = au^2(2b) - (2b)u^2a \in U$. Since $2S = S$ this gives rise to $au^2b - bu^2a \in U$ for all $a \in K$, $b \in S$, $u \in U$, which is Lemma 4.

LEMMA 5. *If $a \in K$, $u \in U$ then $au^4a \in U$.*

Proof. By Lemma 2, $au^2 - u^2a \in U$. Thus $2(au^2 - u^2a)^2 \in U$, and so $4(au^2 - u^2a)^2 = ((2a)u^2 - u^2(2a))^2 \in U$. Since $2K = K$, (as is easily seen) we obtain that $(au^2 - u^2a)^2 \in U$ for all $a \in K$, $u \in U$. But

$$(au^2 - u^2a)^2 = \{(au^2a)u^2 + u^2(au^2a)\} - u^2a^2u^2 - au^4a.$$

Since $au^2a \in S$, as we saw in the proof of Lemma 4, $(au^2a)u^2 + u^2(au^2a) \in U$. Also

$$2u^2a^2u^2 = \{(u^2a^2 + a^2u^2)u^2 + u^2(u^2a^2 + a^2u^2)\} - \{u^4a^2 + a^2u^4\},$$

and since $a^2 \in S$, both $\{\ \}$ are in U, so $2u^2a^2u^2 \in U$, from which, as before, $u^2a^2u^2 \in U$. Thus we obtain that $au^4a \in U$.

Linearizing the lemma we have

COROLLARY. *If $a, b \in K$ then for $u \in U$, $au^4b + bu^4a \in U$.*

We are now able to prove

THEOREM 6. *If $u \in U, r, t \in A$ then $ru^4t + t^*u^4r^* \in U$.*

Proof. Clearly $r = r_0 + r_1$ where $r_0 \in S$, $r_1 \in K$ ($r_0 = \frac{1}{2}(r^* + r)$, $r_1 = \frac{1}{2}(r - r^*)$), and $t = t_0 + t_1$ where $t_0 \in S$, $t_1 \in K$. Thus

$$ru^4t + t^*u^4r^* = (r_0 + r_1)u^4(t_0 + t_1) + (r_0 - r_1)u^4(t_0 - t_1) = (r_0u^4t_0 + t_0u^4r_0) + (r_0u^4t_1 - t_1u^4r_0) + (r_1u^4t_0 - t_0u^4r_1) + (r_1u^4t_1 + t_1u^4r_1).$$

By the corollary to Lemma 5, since $r_1, t_1 \in K$, $(r_1u^4t_1 + t_1u^4r_1) \in U$. Since $2u^2 \in U$, $4u^4 \in U$, and since $2S = S$, by Lemma 3, $(r_0u^4t_0 + t_0u^4r_0) \in U$. Since $2u^2 \in U$, and since $2S = S$, by Lemma 4, $(r_0u^4t_1 - t_1u^4r_0) \in U$ and likewise $(r_1u^4t_0 - t_0u^4r_1) \in U$. Thus each component piece in the expression for $ru^4t + t^*u^4r^*$ is in U, so $ru^4t + t^*u^4r^* \in U$ for all $r, t \in A$.

THEOREM 7. *If $U \neq S$ then $u \in U$ implies that $u^4 = 0$.*

Proof. Suppose that $u^4 \neq 0$ for some $u \in U$. Since A is a simple ring, $Au^4A = A$. Thus if $y \in A$, $y = \sum r_iu^4t_i$ where $r_i, t_i \in A$. Hence $y^* = \sum t_i^*u^4r_i^*$. But then $y + y^* = \sum (r_iu^4t_i + t_i^*u^4r_i^*)$ and so in U since each $r_iu^4t_i + t_i^*u^4r_i^*$

is in U by Theorem 6. That is $y+y^* \varepsilon U$ for every $y \varepsilon A$. Since every element of S is so representable we obtain $U=S$, a contradiction. In this way we are forced to $u^4=0$ for all $u \varepsilon U$.

We are now able to prove the principal theorem of this section.

THEOREM 8. *If A is a simple ring of characteristic $\neq 2$ then S is a simple Jordan ring.*

Proof. Suppose U is a Jordan ideal of S and that $U \neq S$. If $r \varepsilon A$, $u \varepsilon U$ and $r=r_0+r_1$, $r_0 \varepsilon S$, $r_1 \varepsilon K$ then

$$u^2r+r^*u^2=u^2(r_0+r_1)+(r_0-r_1)u^2=(u^2r_0+r_0u^2)+(u^2r_1-r_1u^2),$$

and so is in U by Lemma 2. Consequently by the previous theorem, $(u^2r+r^*u^2)^4=0$. Since $u^4=0$, multiplying $(u^2r+r^*u^2)^4=0$ by u^2 from the right and r from the left we have $r(u^2r+r^*u^2)^4u^2=0$; simplifying this we obtain $(ru^2)^5=0$. But then Au^2 is a left-ideal all of whose elements are nilpotent of index 5; by a theorem of Levitzki [9], Au^2 must be locally nilpotent. If $Au^2 \neq 0$, by another theorem of Levitzki [10], A must possess a non-zero locally nilpotent two-sided ideal. The simplicity of A then forces A to be locally nilpotent; this is impossible in a simple ring [3]. Thus $Au^2=(0)$ and because A is simple, $u^2=0$ results. That is, $u^2=0$ for every $u \varepsilon U$. Therefore for every $u, v \varepsilon U$, $uv+vu=(u+v)^2-u^2-v^2=0$. If $x \varepsilon S$, $v=xu+ux \varepsilon U$, and so $u(xu+ux)+(xu+ux)u=0$. This reduces to $2uxu=0$ since $u^2=0$. A has characteristic $\neq 2$, thus $uxu=0$ for all $x \varepsilon S$. If $a \varepsilon K$, $aua \varepsilon S$, so by the above $uauau=u(aua)u=0$. If r is any element of A, $r=r_0+r_1$, where $r_0 \varepsilon S$, $r_1 \varepsilon K$; whence

$$ururu=u(r_0+r_1)u(r_0+r_1)u=ur_1ur_1u=0.$$

In this way Au is a left-ideal in which every element is nilpotent of index 3. Using the argument as above we are led to $Au=(0)$, and so $u=0$. That is $U=(0)$. Thus the only proper Jordan ideal of S is (0), consequently S is a simple Jordan ring.

The following theorem is a generalization to the case of an arbitrary simple ring of a result Dieudonné proved for division rings [2]. We shall need it for use later in this paper when we study the Lie ideal structure of K.

We first define: If B is a subset of A then $\overline{B}$ is the subring of A generated by B.

THEOREM 9. *If A is a simple ring of characteristic $\neq 2$ and if Z, the center of A, is (0) or if A is more than 4-dimensional over Z, then $\overline{S}=A$.*

Proof. By its very definition $\bar{S}$ is a subring of A. We claim that $\bar{S}$ is, in addition, a Lie ideal of A. For, if $a \in \bar{S}$, $s \in S$, clearly as $-sa \in \bar{S}$. On the other hand, if $b, a \in S$, $k \in S$, then $abk - kab = a(bk - kb) + (ak - ka)b$ and so is in $\bar{S}$ since $[K, S] \subset S$; continuing in this way, we have $ak - ka \in \bar{S}$ for all $a \in \bar{S}$, $k \in K$. Since every $r \in A$ can be written as $r = s + k$ with $s \in S$ and $k \in K$, we obtain that $ar - ra \in \bar{S}$ for all $a \in \bar{S}$ and all $r \in A$. Thus $\bar{S}$ is a Lie ideal of A. By Theorem 3 of [3] either $\bar{S} = A$ or $\bar{S} \subset Z$, the center of A. Suppose $\bar{S} \neq A$; then the other possibility, $\bar{S} \subset Z$, must prevail. In particular $S \subset Z$. Given $r \in A$, $r = k + \lambda$ where $\lambda \in S \subset Z$, $k \in K$. Thus $(r - \lambda)^2 = k^2 \in Z$ since $k^2 \in S$. Therefore every element of A satisfies a quadratic equation over Z. In this discussion $Z \neq (0)$ (otherwise $\bar{S} \subset Z$ is nonsensical), so A is a primitive ring, in which every element satisfies an equation of degree 2 over the center. By the work of Jacobson [6] this implies that A is at most 4-dimensional over Z. The proof of Theorem 9 is thereby completed.

3. The Lie ideals of K. We now turn our attention to the problem of finding the possible Lie ideals of K. Let U be a Lie ideal of K; our aim is to show that either $U \subset Z$ or $U \supset [K, K]$. Our starting point in these considerations is

LEMMA 10. *If $u \in U$, $x \in S$ then $u^2x - xu^2 \in U$.*

Proof. Since $u \in U \subset K$, $x \in S$, it follows that $xu + ux \in K$. Consequently $u^2x - xu^2 = (xu + ux)u - u(xu + ux) \in U$ since U is a Lie ideal of K.

In the case of Jordan ideals we had that u in the ideal implies that $2u^2$ is also in that ideal. In the Lie case we have no such closure under squaring, but the next lemma provides a substitute, showing that although u^3 need not be in U although $u \in U$, u^3 does have a close relation to U.

LEMMA 11. *If $u \in U$ and $x \in K$ then $u^3x - xu^3 \in U$; that is $[u^3, K] \subset U$.*

Proof. Since $xu + ux \in S$, by Lemma 10, $(xu + ux)u^2 - u^2(xu + ux) \in U$. In other words, $xu^3 - u^3x + uxu^2 - u^2xu \in U$. However, $uxu \in K$, from which it follows that $uxu^2 - u^2xu = uxu \cdot u - u \cdot uxu \in U$. By the above we are left with $xu^3 - u^3x$ completing the proof.

For U a Lie ideal of K, we define $T(U)$ by $T(U) = \{x \in K \mid [x, K] \subset U\}$. We note some properties of $T(U)$ in

LEMMA 12. *$T(U)$ is a Lie ideal of K; moreover $U \subset T(U)$.*

Proof. That $U \subset T(U)$ is, of course, nothing more than the definition of a Lie ideal of K. It is also clear that $T(U)$ is an additive subgroup of U.

To prove the lemma there remains but to show that $[T(U),K]\subset T(U)$. Since $[T(U),K]\subset U$ by definition, and since $U\subset T(U)$, this fact follows easily also.

Although $u\in U$ need not imply that $u^3\in U$ we note

LEMMA 13. *$t\in T(U)$ implies that $t^3\in T(U)$.*

Proof. Suppose that $t\in T(U)$, $x\in K$. We must show that $t^3x-xt^3\in U$. Now if $s\in S$, $t^2s-st^2=t(ts+st)-(ts-st)t$ is in U since $ts+st\in K$. In particular, for $s=tx+xt$,

$$U\ni t^2(tx+xt)-(tx+xt)t^2=t^3x-xt^3+t\cdot txt-txt\cdot t.$$

Since $txt\in K$, $txt\cdot t-t\cdot txt\in U$; whence $t^3x-xt^3\in U$ and $t^3\in T(U)$.

Although we do not need the concept in this paper, we feel that it will prove useful in further considerations on the Lie structure of rings with involutions, so we define:

U is a *strong Lie ideal* of K if $u\in U$ implies that $u^3\in U$.

We note the following corollary to Lemma 13:

COROLLARY. *If $U\supset[K,K]$ is a Lie ideal of K then U can be imbedded in a proper, strong Lie ideal of K.*

For if $U\supset[K,K]$, $T(U)\neq K$. Since, by Lemma 13, $T(U)$ is a strong Lie ideal of K and $U\subset T(U)$ the corollary follows.

Let $K\circ K$ be the additive subgroup generated by all $xy+yx$ where x and y range over K. Clearly $K\circ K\subset S$.

LEMMA 14. $S=[K,S]+K\circ K$.

Proof. As noted above, $K\circ K\subset S$. It is equally trivial that $[K,S]\subset S$. We claim that $[K,S]+K\circ K$ is a Jordan ideal of S. It is obviously an additive subgroup of S. Suppose now that $a,b\in K$ and $s\in S$. Thus

$$(ab+ba)s+s(ab+ba)$$
$$=\{a(bs-sb)-(bs-sb)a\}+\{(as+sa)b+b(as+sa)\}.$$

The first { } is in $[K,S]$ since $bs-sb\in S$; since $as+sa\in K$, the second { } is in $K\circ K$. Consequently $(K\circ K)\circ S\subset[K,S]+K\circ K$.

On the other hand, if $a\in K$, $s,t\in S$ then

$$(as-sa)t+t(as-sa)$$
$$=\{a(st-ts)+(st-ts)a\}+\{(at+ta)s-s(at+ta)\}.$$

Since $st - ts \in K$, the first $\{\ \}$ is in $K \circ K$; since $at + ta \in K$, the second $\{\ \}$ is in $[K, S]$. Thus $[K, S] \circ S \subset [K, S] + K \circ K$.

In other words, $[K, S] + K \circ K$ is a Jordan ideal of S. By Theorem 8 it follows that $S = [K, S] + K \circ K$, (the desired result) or $[K, S] + K \circ K = (0)$. We wish to rule out the second possibility. If $[K, S] + K \circ K = (0)$, then in particular, $k \in K$ implies that $k^2 = 0$. Also, $k \in K$ and $s \in S$ implies that $ks = sk$. Thus $0 = k^2 s = ksk$; that is $kSk = (0)$. If, on the other hand, $a \in K$, then $ka + ak = 0$ since it is in $K \circ K = (0)$; multiplying this on the right by k we obtain $kak = 0$; hence $kKk = (0)$. Since $A = K + S$, $kAk = k(K + S)k = kKk + kSk = (0)$. kA is, in this way, a nilpotent right ideal, which is impossible in a simple ring unless $k = 0$. Since $K \neq (0)$ (otherwise A is a commutative field), $[K, S] + K \circ K = (0)$ is not a possibility, and so it is equal to the only other possibility, namely S.

We recall that $\bar{K}$ is the subring of A generated by K. Although the next theorem is of some independent interest, it is the essential key to all the results that follow.

THEOREM 15. *If A has a trivial center or if A is more than 4-dimensional over its center then $\bar{K} = A$.*

Proof. By definition, $K \subset \bar{K}$. We now consider those cases for which $S \subset \bar{K}$ will be provable.

Let $a, b \in K$, and let $s \in S$. Thus $sa + as \in K$, and so $(sa + as)b \in \bar{K}$. That is, $sab + asb \in \bar{K}$. Similarly $sba + bsa \in \bar{K}$. Subtracting these two we obtain that $s(ab - ba) + asb - bsa \in \bar{K}$. However, $asb - bsa = asb - b^* s^* a^* = asb - (asb)^*$ and thus is in K, and so in $\bar{K}$. Consequently we have demonstrated that $s(ab - ba) \in \bar{K}$. Rephrasing this, $S(ab - ba) \subset \bar{K}$. However, since $ab - ba \in K$, $K(ab - ba) \subset \bar{K}$, and since $A = K + S$, we obtain that $A(ab - ba) \subset K$ for all $a, b \in K$. Similarly, if $c, d \in K$, $(cd - dc)A \subset \bar{K}$. Since $\bar{K}$ is a subring of A, and since both $A(ab - ba)$ and $(cd - dc)A$ are contained in $\bar{K}$, $A(ab - ba)(cd - dc)A \subset \bar{K}$. A is a simple ring and $A(ab - ba)(cd - dc)A$ is a two-sided ideal of A, so either $A(ab - ba)(cd - dc)A = A$, in which case $A \subset \bar{K}$, the desired result, or $A(ab - ba)(cd - dc)A = (0)$ for all $a, b, c, d \in K$. We consider when the second possibility can hold. In that case, by the simplicity of A, $(ab - ba)(cd - dc) = 0$ for all $a, b, c, d \in K$. Since $K(cd - dc)A \subset \bar{K}$, the same argument as used above leads to $(ab - ba)\bar{K}(cd - dc) = (0)$ for all $a, b, c, d \in K$. Suppose now that $s \in S$, $c, d \in K$. Then

$$s(cd + dc) - (cd + dc)s$$
$$= \{(sc + cs)d - d(sc + cs)\} + \{(ds + sd)c - c(ds + sd)\}.$$

Since $sc + cs \in K$, the first $\{\ \}$ is in $[K, K]$; similarly the second $\{\ \}$ is in $[K, K]$. Thus the left-hand side is in $[K, K]$. But then, since we know from the argument above that $(ab - ba)[K, K] = (0)$ for $a, b \in K$, we have that $(ab - ba)\{s(cd + dc) - (cd + dc)s\} = (0)$ for all $a, b, c, d \in K$ and $s \in S$. If we further suppose that $d = ef - fe$ where $e, f \in K$ then $(ab - ba)dc = 0$ and $(ab - ba)cd = 0$ since it is contained in $(ab - ba)\bar{K}d = (0)$; and so $(ab - ba)(cd + dc) = 0$. Thus $(ab - ba)\{(cd + dc)s - s(cd + dc)\} = 0$ reduces to $(ab - ba)s(cd + dc) = 0$ for all $s \in S$, when $d = ef - fe$. That is, $(ab - ba)S(c(ef - fe) + (ef - fe)c) = (0)$ for all $a, b, c, e, f \in K$. Since, in addition, $(ab - ba)K(c(ef - fe) + (ef - fe)c) = (0)$ (since it is contained in $(ab - ba)\bar{K}(ef - fe) = (0)$) and since $A = K + S$, we obtain $(ab - ba)A(c(ef - fe) + (ef - fe)c) = (0)$ for all $a, b, c, e, f \in K$. We wish to show that $ab = ba$ for all $a, b \in K$. If not, that is, if $ab - ba \neq 0$ for some $a, b \in K$, then by the simplicity of A, $c(ef - fe) + (ef - fe)c = 0$ for all c, e, f in K. If $s \in S$ then $c = (ef - fe)s + s(ef - fe)$ is in K and so $c(ef - fe) + (ef - fe)c = 0$ implies, since $(ef - fe)^2 = 0$, that $2(ef - fe)s(ef - fe) = 0$; that is, $(ef - fe)S(ef - fe) = (0)$. Since, from before, $(ef - fe)K(ef - fe) = (0)$ we are led to $(ef - fe)A(ef - fe) = (0)$, which forces $ef - fe = 0$ by the simplicity of A. So, $\bar{K} \neq A$ has resulted in $ab = ba$ for all $a, b \in K$. In particular, if $a \in K$, a^2 commutes with all elements of K. Now if $a \in K$, $s \in S$, then $as + sa \in K$, and so $a(as + sa) = (as + sa)a$; in simplifying this says that $a^2s = sa^2$ for all $s \in S$. Since a^2 commutes with all elements of K and of S, and since $K + S = A$, a^2 commutes with all elements of A. Consequently $a^2 \in Z$, the center of A, for all $a \in K$. Linearizing this we obtain that $ab + ba$ is in Z for all $a, b \in K$, and so $K \circ K \subset Z$. Since $ab = ba$, we have that $2ab = ab + ba \in Z$. Thus $a, b \in K$ implies that $ab \in Z$. Since $Z = (0)$ or Z is a field, $ab = 0$ or a has an inverse in A for all $a, b \in K$. If a has an inverse for some $a \in K$ then so does b for every $b \neq 0 \in K$, for $ab \neq 0 \in Z$ and so has an inverse, whence b has an inverse in A. We rule out the possibility that no $a \in K$ has an inverse in A. For then, since $a^2 \in Z$, $a^2 = 0$ results; also, since $ab \in Z$, $ab = 0$ for all $b \in K$. Since $b = as + sa \in K$, for $s \in S$, $a(as + sa) = 0$, and so $aSa = (0)$; since $aKa = a^2K = (0)$, we obtain $aAa = (0)$, and so $a = 0$ because A is a simple ring. Thus we must assume that every $a \neq 0 \in K$ has an inverse in A. If $b \neq 0 \in K$, $ab = \lambda \in Z$, $\lambda \neq 0$, whence $a \cdot ab = \lambda a$, and since $a^2 = \mu \neq 0 \in Z$, we are left with $b = \lambda' a$ where $\lambda' \in Z$. Thus $K = \{\lambda a\}$ for appropriate λ''s in Z.

If $s \in S$ then $(as - sa)^2 = a \cdot sas + sas \cdot a - as^2a - sa^2a$, and since $sas \in K$, $a \cdot sas + sas \cdot a \in Z$; also $as^2a + sa^2s = as^2a + a^2s^2$ (since $a^2 \in Z$) $= a(s^2a + as^2) \in Z$ since $s^2a + as^2 \in K$. Thus $(as - sa)^2 \in Z$ for all $s \in S$. Since $S = [K, S]$

$+K \circ K$, and since $K = \{\lambda a\}$, and since $K \circ K \subset Z$, we have that every element s of S can be written as $s = \lambda(at - ta) + \mu$ for $\lambda, \mu \in Z$ and $t \in S$. Since every $r \in A$ can be written as $r = s + k$ we have that $r = \lambda(at + ta) + \mu + \alpha a$ for $\lambda, \mu, \alpha \in Z$, $t \in S$. Thus

$$(r - \mu)^2 = \lambda^2(at - ta)^2 + \alpha^2 a^2 + \alpha\lambda\{a(at - ta) + (at - ta)a\}.$$

Now, $(at - ta)^2 \in Z$, $a^2 \in Z$ and $a(at - ta) + (at - ta)a = a^2t - ta^2 = 0$, so $(r - \mu)^2 \in Z$. Thus every $r \in A$ satisfies a quadratic equation over Z. As in the argument used in proving Theorem 9, A is at most 4-dimensional over Z. Theorem 15 is now completely proved.

It will be useful to show

LEMMA 16.[1] *If $u, v \in U$ then $uvu \in T(U)$ and $u^2v + vu^2 \in T(U)$.*

Proof. To prove that $uvu \in T(U)$ we need but verify that $[uvu, K] \subset U$. Let $x \in K$. Thus

$$uvux - xuvu = \{u(vux + xuv) - (vux + xuv)u\} + \{vuxu - uxuv\}.$$

But $vux + xuv = vux - (vux)^* \in K$, so the first $\{\ \}$ is in U since U is a Lie ideal of K. Since $uxu \in K$ and $v \in U$, the second $\{\ \}$ is also in U. Thus $uvux - xuvu \in U$, and so $uvu \in T(U)$. Now

$$u^2v + vu^2 = \{u(uv - vu) - (uv - vu)u\} + 2uvu,$$

and since $uvu \in T(U)$ and since $u(uv - vu) - (uv - vu)u \in U \subset T(U)$, $u^2v + vu^2$ also is in $T(U)$, proving the lemma.

If U is a Lie ideal of K we define $B(U)$ by:

$$B(U) = \{x \in S \mid xu + ux \in T(U) \text{ for all } u \in U\}.$$

By Lemma 16, $u \in U$ implies that $u^2 \in B(U)$.

LEMMA 17. *Let $x \in B(U)$; then $(xu - ux)y + y(xu - ux) \in T(U)$ for all $y \in K$.*

Proof.

$$(xu - ux)y + y(xu - ux)$$
$$= \{x(uy - yu) + (uy - yu)x\} + \{(xy + yx)u - u(xy + yx)\}.$$

Since $uy - yu \in U$, and since $x \in B(U)$, the first $\{\ \}$ on the right-hand side is in $T(U)$. Since $xy + yx \in K$, and since $u \in U$, a Lie ideal of K, the second $\{\ \}$ is in U, and so is in $T(U)$; this proves the lemma.

[1] The proof is patterned after a suggestion of Willard E. Baxter.

The lemma motivates the following definition: for U a Lie ideal of K,

$$C(U) = \{x \in S \mid xy + yx \in T(U) \text{ for all } y \in K\}.$$

By Lemma 17, $[U, B(U)] \subset C(U)$.

Let $u, v \in U$. Thus $u^2 \in B(U)$. Hence $u^2v - vu^2 \in C(U)$. Therefore $(u^2v - vu^2)k + k(u^2v - vu^2) \in T(U)$ for all $k \in K$. On the other hand, if $s \in S$,

$$(u^2v - vu^2)s - s(u^2v - vu^2)$$
$$= \{u^2(vs - sv) - (vs - sv)u^2\} + \{(u^2s - su^2)v - v(u^2s - su^2)\}.$$

Since $vs - sv \in S$, the first $\{\ \}$ is in U by Lemma 10. Since $u^2s - su^2 \in K$, the second $\{\ \}$ is also in U. Thus the right-hand side is in U, consequently it is in $T(U)$. Thus $(u^2v - vu^2)s - s(u^2v - vu^2) \in T(U)$ for all $s \in S$. Given any element $r \in A$, $r = r_0 + r_1$ where $r_0 \in S$ and $r_1 \in K$. Computing $(u^2v - vu^2)r - ((u^2v - vu^2)r)^*$ in terms of r_0 and r_1 and using the above discussion we see that $(u^2v - vu^2)r - ((u^2v - vu^2)r)^*$ is in $T(U)$ for all $r \in A$. We summarize this in

THEOREM 18. *Let U be a Lie ideal of K, and let $u, v \in U$, $r \in A$; then*

$$(u^2v - vu^2)r - ((u^2v - vu^2)r)^* \in T(U).$$

We now define for U a Lie ideal of K, $G(U)$ by

$$G(U) = \{g \in A \mid gr - r^*g^* \in T(U) \text{ for all } r \in A\}.$$

By Theorem 18, $u^2v - vu^2 \in G(U)$ for all $u, v \in U$.

THEOREM 19. *Let A be a simple ring of characteristic $\neq 2$ and suppose that either Z, the center of A is (0) or that A is more than 4-dimensional over Z. If U is a Lie ideal of K and if $U \supset [K, K]$ then $G(U) = (0)$.*

Proof. Let $g \neq 0 \in G(U)$; thus for any $r \in A$

$$gr - r^*g^* \in T(U). \tag{1}$$

Thus, if $k \in K$, since $T(U)$ is a Lie ideal of K,

$$(gr - r^*g^*)k - k(gr - r^*g^*) \in T(U).$$

However,

$$(gr - r^*g^*)k - k(gr - r^*g^*) = g(rk) + (kr^*)g^* - r^*g^*k - kgr \in T(U).$$

Since $grk + kr^*g^* = g(rk) - (rk)^*g^*$, it is in $T(U)$ by (1). Thus we obtain

$$r^*g^*k + kgr \in T(U) \text{ for all } r \in A,\ k \in K. \tag{2}$$

Let $k_1, k_2 \in K$. Since $T(U)$ is a Lie ideal of K, by (2) we have that

$$(r^*g^*k_1 + k_1gr)k_2 - k_2(r^*g^*k_1 + k_1gr) \in T(U).$$

That is,

$$k_1grk_2 - k_2r^*g^*k_1 + (rk_2)^*g^*k_1 - k_2k_1gr \in T(U).$$

However,

$$k_1grk_2 - k_2r^*g^*k_1 = k_1g(rk_2) + (rk_2)^*g^*k_1$$

and so is in $T(U)$ by (2). We thus are led to $r^*g^*k_1k_2 - (k_1k_2)^*gr \in T(U)$. Continuing in this way we obtain $r^*g^*\bar{k} - (\bar{k})^*gr \in T(U)$ for all $\bar{k} \in \bar{K}$ and all $r \in A$. By Theorem 15, $\bar{K} = A$ if the center of A is trivial or if A is more than 4-dimensional over Z, so in these cases we have that $r^*g^*t^* - tgr \in T(U)$ for all $t, r \in A$. Since A is simple, and since $g \neq 0$, $A = AgA$. Thus, given $y \in A$, $y = \sum r_igt_i$. Then $y^* = \sum t_i^*g^*r_i^*$, and so $y - y^* = \sum(r_igt_i - t_i^*g^*r_i^*)$. Since each $r_igt_i - t_i^*g^*r_i^* \in T(U)$ by the discussion above, we obtain that $y - y^* \in T(U)$ for all $y \in A$. However, every element in K has such a representation in the form $y - y^*$; we thus obtain $T(U) = K$. By the definition of $T(U)$ this is equivalent with $U \supset [K, K]$. Since we assumed $D \not\supset [K, K]$ we are led to a contradiction, and so $G(U) = (0)$.

We assume henceforth that $Z = (0)$ or that A is more than 4-dimensional over Z.

Since for U a Lie ideal of K, $u^2v - vu^2 \in G(U)$ for all $u, v \in U$, and since, by Theorem 19, if $U \not\supset [K, K]$, $G(U) = (0)$, we have

THEOREM 20. *If $U \not\supset [K, K]$ is a Lie ideal of K, and if $u, v \in U$, then $u^2v = vu^2$.*

We now prove

THEOREM 21. *Suppose U is a Lie ideal of K in which $u^2 = 0$ for every $u \in U$. Then $U = (0)$.*

Proof. We suppose that $u^2 = 0$ for every $u \in U$. Let $u \in U$, $k \in K$. Thus $2uku = (uk - ku)u - u(uk - ku) \in U$. Since $2K = K$, we obtain $uku \in U$ for all $k \in K$ and all $u \in U$. If $u, v \in U$ then $uv + vu = (u + v)^2 - u^2 - v^2 = 0$, so left-multiplying this by v we obtain $vuv = 0$ for all $v, u \in U$. Since $uku \in U$, we also have $v(uku)v = 0$. That is, $vukuv = 0$ for all $u, v \in U$ and $k \in K$; since $uv = -vu$ we can obtain $vukvu = 0$ for all $v, u \in U$ and all $k \in K$. Let $w \in U$ and $s \in S$. Thus $sw + ws \in K$, whence $vu(sw + ws)vu = 0$. Left-multiplying this by w we obtain $wvu(sw + ws)vu = 0$. However, since

$w \in U$, $wvu = -vwu = vuw$, thus $wvuwsvu = vuw^2svu = 0$, and since $0 = wvu(sw + ws)vu$, this reduces to $(wvu)s(wvu) = 0$; in other form this says $wvuSwvu = (0)$. Since we have already established that $wvuKwvu = (0)$, and since $A = K + S$, these combine to yield $wvuAwvu = (0)$. The simplicity of A then leads to $wvu = 0$ for all $w, v, u \in U$. If $k \in K$, let $w = ku - uk \in U$; hence $(ku - uk)vu = 0$, and since $uvu = 0$, we arrive at $ukvu = 0$ for all $k \in K$. Put $k = sv + vs$ where $s \in S$. Then $0 = u(sv + vs)vu = uvsvu = uvsuv$ since $v^2 = 0$, $uv = -vu$. That is, $uvSuv = (0)$. As we already have established that $uvKuv = (0)$, we reach $uvAuv = (0)$, and so $uv = 0$ for all $u, v \in U$. Put $v = uk - ku$ where $k \in K$. Since $u^2 = 0$, $uv = 0$ yields that $uku = 0$ for all $k \in K$. If $s \in S$, $sus \in K$ and so $u(sus)u = 0$, by the above. If $r \in A$, then $r = s + k$ when $s \in S$, $k \in K$ and so $ururu = u(s + k)u(s + k)u = ususu = 0$ since $uku = 0$. Thus uA is a nil right-ideal of A in which each element is nilpotent of index 3. As we have previously shown in this paper this is impossible in a simple ring unless $uA = (0)$; but then $u = 0$. Thus we have shown that $U = (0)$, proving Theorem 21.

THEOREM 22. *Let U be a Lie ideal of K and suppose that $U \not\subset [K, K]$. Then $u \in U$ implies that $u^2 \in Z$, the center of A.*

Proof. Since $u^2v = vu^2$ for all u, v in U, by Theorem 20, u^2 is in the center of $\bar{U}$, the subring generated by U. Now, $u^2s - su^2 \in U \subset \bar{U}$ by Lemma 10 for all $s \in S$. Also $u^2k - ku^2 = (uk - ku)u + u(uk - ku)$ and since $u \in U$, $uk - ku \in U$, each of $u(uk - ku)$ and $(uk - ku)u$ is in $\bar{U}$ for $k \in K$; thus $u^2k - ku^2 \in \bar{U}$. But then $u^2a - au^2 \in \bar{U}$ for all $a \in A$. Since u^2 is in the center of $\bar{U}$, $u^2(u^2a - au^2) = (u^2a - au^2)u^2$ for all $a \in A$, $u \in U$. The theorem will thus be proved when we prove

SUBLEMMA. *Let A be a simple ring of characteristic $\neq 2$. Suppose $t \in A$ is such that $t(ta - at) = (ta - at)t$ for all $a \in A$. Then $t \in Z$.*

This sublemma has some independent interest. To prove the sublemma we proceed as follows.

We know that $t(tr - rt) = (tr - rt)t$ for all $r \in A$. Let $p \in A$. Thus $t(trp - rpt) = (trp - rpt)t$. But $trp - rpt = (tr - rt)p + r(tp - pt)$. Thus $t(trp - rpt) = t\{(tr - rt)p + r(tp - pt)\} = (tr - rt)tp + tr(tp - pt)$ since $tr - rt$ commutes with t. Similarly $(trp - rpt)t = (tr - rt)pt + rt(tp - pt)$. Equating the two, transposing and simplifying we arrive at $2(tr - rt)(tp - pt) = 0$, and since the characteristic of A is not 2, we have that $(tr - rt)(tp - pt) = 0$ for all $r, p \in A$. In particular, for $p = ar$,

$$(1) \qquad (tr - rt)(tar - art) = 0 \qquad \text{for all } a, r \in A.$$

Since $tar - art = (ta - at)r + a(tr - rt)$ and since $(tr - rt)(ta - at) = 0$, (1) yields that $(tr - rt)a(tr - rt) = 0$; that is, $(tr - rt)A(tr - rt) = (0)$. This is impossible in a simple ring unless $tr - rt = 0$. Thus $t \in Z$ and the sublemma is established. Since $u^2(u^2a - au^2) = (u^2a - au^2)u^2$ in the theorem, $u^2 \in Z$ follows, and the theorem is proved.

COROLLARY. *Let U be a Lie ideal of K and suppose that $U \not\supset [K, K]$. Then $u, v \in U$ implies that $uv + vu \in Z$.*

We now are able to dispose of the situation in which $Z = (0)$. Indeed

THEOREM 23. *Let A be a simple ring of characteristic $\neq 2$ whose center $Z = (0)$. If $U \neq (0)$ is a Lie ideal of K then $U \supset [K, K]$.*

Proof. If $U \not\supset [K, K]$ then by Theorem 22 $u \in U$ implies that $u^2 \in Z = (0)$. Consequently $u^2 = 0$ for all $u \in U$. By Theorem 21 this results in $U = (0)$.

Having, in this way, settled the case $Z = (0)$ *we henceforth assume that* $Z \neq (0)$.

The * of A induces an automorphism on Z. Two possibilities now confront us, namely

(1) $\lambda^* = \lambda$ for all $\lambda \in Z$, an involution of the first kind,

(2) $\mu^* \neq \mu$ for some $\mu \in Z$, an involution of the second kind.

In the case of an involution of the second kind, $\lambda = \mu^* - \mu \neq 0$ is in Z and is such that $\lambda^* = -\lambda$. Thus (2) is equivalent to

(2′) $\mu^* = -\mu$ for some $\mu \neq 0 \in Z$.

Our discussion first turns to the case (2′) in which there is a skew element in the center of A.

Let $\mu \in Z$, $\mu^* = -\mu \neq 0$. Let U be a Lie ideal of K and we further suppose that $U \not\supset [K, K]$.

If $s \in S$ then $\mu s \in K$, hence for all $u \in U$, $u(\mu s) - (\mu s)u \in U$; that is $\mu(us - su) \in U$. If $v \in U$, by the corollary to Theorem 22,

$$\mu(us - su)v + v(\mu(us - su)) \in Z. \quad \text{That is } \mu((us - su)v + v(us - su)) \in Z.$$

As a consequence, since $\mu \neq 0 \in Z$, $(us - su)v + v(us - su) \in U$ for all $s \in S$ and all $u, v \in U$. In particular, if $k \in K$, $k^2 \in S$, whence

$$z = (uk^2 - k^2u)v + v(uk^2 - k^2u)$$

is in Z. Since $uk^2 - k^2u = (uk - ku)k + k(uk - ku)$, it is readily verified that

$$z = \{(uk-ku)(vk-kv) - (vk-kv)(uk-ku)\}$$
$$-\{((uk-ku)v + v(uk-ku))k + k((uk-ku)v + v(uk-ku))\}.$$

Since both v and $uk-ku$ are in U, $(uk-ku)v + v(uk-ku) = \lambda u, v) \in Z$. Also, since $uk-ku$ and $vk-kv$ are in U, the first $\{\ \}$ on the right-hand side is in U. Thus $2\lambda(u,v)k = u_1 + z_1$ where $u_1 \in U$, $z_1 \in Z \cap K$, $(z_1 = -z)$. If $\lambda(u,v) \neq 0$ for some $u, v \in U$, since $\lambda(u,v)$ is also in S we would have that $k \in U + Z$. This is precisely what we need, so we claim next that given $k \in K$, $\lambda(u,v) \neq 0$ for some $u, v \in U$. That is, we claim that it is impossible that $(uk-ku)v + v(uk-ku) = 0$ for all $u, v \in U$ where $k \notin U + Z$, $k \in K$. For, if so, let $w = uk-ku$. By assumption $wv + vw = 0$ for all $v \in U$. Thus, if $v, x \in U$, since $wv + vw = wx + xw = 0$, w commutes with vx and with xv, and so with $vx - xv$. Since $vx - xv \in U$, $w(vx-xv) + (vx-xv)w = 0$; that is, $2w(vx-xv) = 0$, whence $w(vx-xv) = 0$. We have thus obtained $w[U,U] = (0)$. Since $[U,U] \subset U$ is a Lie ideal of K and does not contain $[K,K]$, if $[U,U] \neq (0)$ by Theorem 21 there is an element $t \in [U,U]$ with $t^2 \neq 0 \in Z$. Since $wt = 0$ we obtain $w = 0$. If, on the other hand, $[U,U] = (0)$, then given $u, v \in U$, $uv = vu$. Since $uv + vu \in Z$ we obtain that $2uv \in Z$, whence $uv \in Z$. If $u^2 \neq 0$, then $uv \neq 0$ since $u^2 \in Z$ and so has an inverse in A. Since $uv = \lambda \neq 0 \in Z$ and since $u^2 \in Z$, we have that $v = \lambda' u$ for some $\lambda' \in Z$. That is, $U = \{\lambda u\}$ for appropriate λ's in Z. Since $uk' - k'u \in U$ for all $k' \in K$, $uk' - k'u = \lambda u$. Since, however,

$$u(uk' - k'u) + (uk' - k'u) = u^2k' - k'u^2 = 0$$

we see that $2\lambda u^2 = 0$, and so $\lambda = 0$. That is $uk' = k'u$ for all $k' \in K$. But if the dimension of A over Z exceeds 4, $\bar{K} = A$, and since $u \in U$ commutes with every element of K, it does so with every element of $\bar{K}$. That is, $U \subset Z$. This is precisely what we should like to prove. So we assume this possibility is ruled out. This returns us to our only other possibility, namely that $w = uk - ku = 0$ for all $u \in U$ and all $k \notin U + Z$. Thus $uk = ku$ for all $u \in U$ and all $k \notin U + Z$. Suppose that $U + Z \cap K \neq K$. Then there is an element $k_0 \in K$, $k_0 \notin U + Z \cap K$. Thus $k_0 u = uk_0$ for all $u \in U$. If $v \in U$, then $k_0 + v$ is also not in $U + Z \cap K$, thus $(k_0 + v)u = u(k_0 + v)$. Since $uk_0 = k_0 u$, this leads to $uv = vu$ for all $u, v \in U$. As we saw above, this forces $U \subset Z$. So the assumption $U \not\subset Z$ leads to $K = U + Z \cap K$. But then $[K,K] = [U,U] \subset U$, contradicting that $U \supset [K,K]$. We summarize all this into

THEOREM 24. *Let A be a simple ring of characteristic $\neq 2$ and suppose that $Z \neq (0)$. We further suppose that $\lambda^* = -\lambda \neq 0$ for some*

$\lambda \in Z$. *If A is more than 4-dimensional over Z and if U is a Lie ideal of K then*

(1) *either* $U \subset Z$ *or* (2) $U \supset [K, K]$.

The situation (2′) where $\lambda^* = -\lambda \neq 0$ for some $\lambda \in Z$ is now settled.

We now turn to case (1) in which $\lambda^* = \lambda$ for all $\lambda \in A$. That is, $Z \cap K = (0)$. Our goal in this situation is to prove that any Lie ideal U of K must contain $[K, K]$. With this objective in mind we contend, to begin with, that there exist elements u and v in U so that

(1) $u^2 = \lambda_1 \neq 0 \in Z$, (2) $v^2 = \lambda_2 \neq 0 \in Z$, (3) $uv + vu = 0$.

Since $U \neq (0)$, by Theorem 21 there is an element $u \in U$ so that $u^2 \neq 0$; since $u^2 \in Z$, $u^2 = \lambda_1 \neq 0$, $\lambda_1 \in Z$. Now, if $uk = ku$ for all $k \in K$, by Theorem 15, since $\bar{K} = A$ we would have that $u \in Z$, contradicting $K \cap Z = (0)$. So, for some $k \in K$, $v = uk - ku \neq 0 \in U$. Clearly $uv + vu = 0$. So far we have fulfilled conditions (1) and (3) of our contention. If $(uk - ku)^2 \neq 0$ our contention will have been established in its entirety. So we suppose that $(uk - ku)^2 = 0$ for all $k \in K$. Let $v_1 = uk_1 - k_1u$, $v_2 = uk_2 - k_2u$ where k_1 and k_2 are in K. Thus $uv_1 + v_1u = uv_2 + v_2u = 0$. Hence $u(v_1v_2 - v_2v_1) = (v_1v_2 - v_2v_1)u$. Since $v_1v_2 - v_2v_1 \in U$, $u(v_1v_2 - v_2v_1) + (v_1v_2 - v_2v_1)u \in Z$, that is, $2u(v_1v_2 - v_2v_1) \in Z$, and so $u(v_1v_2 - v_2v_1) \in Z$. However, since $v_1^2 = v_2^2 = 0$ and since $u(v_1 + v_2) + (v_1 + v_2)u = 0$, if our contention were false, $(v_1 + v_2)^2 = 0$, from which $v_1v_2 + v_2v_1 = 0$. But then $(v_1v_2 - v_1v_2)^2 = v_1v_2v_1v_2 + v_2v_1v_2v_1 = 0$. Since $u(v_1v_2 - v_2v_1) \in Z$, if it were not 0, $v_1v_2 - v_2v_1$ would have an inverse in A. Thus $u(v_1v_2 - v_2v_1) = 0$; but since $u^2 = \lambda_1 \neq 0 \in Z$, we obtain $v_1v_2 - v_2v_1 = 0$. Together with the fact established above that $v_1v_2 + v_2v_1 = 0$, this implies that $v_1v_2 = 0$. Another way of stating this is really that if $uv + vu = uw + wu = 0$ then $vw = 0$, when $v, w \in U$. Let

$$w = (uk_2 - k_2u)k_2 - k_2(uk_2 - k_2u); \quad uw + wu = -2(uk_2 - k_2u)^2 = 0.$$

Thus $v_1w = 0$; that is, $(uk_1 - k_1u)\{(uk_2 - k_2u)k_2 - k_2(uk_2 - k_2u)\} = 0$. Since $(uk_1 - k_1u)(uk_2 - k_2u) = 0$ this reduces to $(uk_1 - k_1u)k_2(uk_2 - k_2u) = 0$ for all $k_1, k_2 \in K$. Now

$$u(kk' - k'k) - (kk' - k'k)u$$
$$= (uk - ku)k' - k'(uk - ku) + k(uk' - k'u) - (uk' - k'u)k.$$

Right-multiplying this by $uk - ku$ we obtain, since $kk' - k'k \in K$, that $0 = (uk - ku)k'(uk - ku) - (uk' - k'u)k(uk - ku)$; and since by the above

discussion $(uk'-k'u)k(uk-ku)=0$, we are left with $(uk-ku)K(uk-ku)=(0)$. If $s\in S$, $s(uk-ku)s\in K$, and so $(uk-ku)s(uk-ku)s(uk-ku)=0$. If $r\in A$, $r=s+k'$ where $s\in S$ and $k'\in K$. By the above discussion it follows that $(uk-ku)(s+k')(uk-ku)(s+k')(uk-ku)=0$. So $(uk-ku)A$ is a right-ideal of A each of whose elements is nilpotent of index of nilpotence 3, and as we have seen before, this forces $uk-ku=0$ for all $k\in K$. Since $\bar{K}=A$ this places u in Z, violating $Z\cap K=(0)$. Thus for some $v\in U$ with $uv+vu=0$ it must be that $v^2=\lambda_2\neq 0\in Z$. This proves our contention.

So, we now have elements $u, v\in U$ such that $u^2=\lambda_1\neq 0$, $v^2=\lambda_2\neq 0$, $\lambda_1,\lambda_2,\in Z$ and $uv+vu=0$. A simple calculation verifies that $(uv-vu)^2=-4\lambda_1\lambda_2\neq 0$.

We claim that u, v and $uv-vu$ are linearly independent over Z. For if $w=\lambda_0 u+\lambda_1 u+\lambda_2(uv-vu)=0$, $\lambda_i\in Z$, $uw+wu=0$ yields $2\lambda_0=0$, $vw+wv=0$ yields $2\lambda_1=0$ and $(uv-vu)w+w(uv-vu)=0$ yields that $2\lambda_2=0$. Hence $\lambda_0=\lambda_1=\lambda_2=0$, and so u, v, $uv-vu$ are linearly independent over Z.

Suppose now that we could find an $x\in U$ so that u, v, $uv-vu$ and x are linearly independent over Z. Suppose that $ux+xu=\lambda$, $vx+xv=\mu$ where $\lambda,\mu\in Z$. Consider $x'=x+\alpha u+\beta v$ where α and β are in Z. Now, $x'u+ux'=\lambda+2\alpha\lambda_1$, $x'v+vx'=\mu+2\beta\lambda_2$. Since $\lambda_1\neq 0$, $\lambda_2\neq 0$ we can solve these for α,β to force $x'u+ux'=x'v+vx'=0$. Note that x' is linearly independent of u, v and $uv-vu$ and that it is in U, (since $Z\subset S$). We drop the $'$. Since $xu+ux=xv+vx=0$, we have that $x(uv-vu)=(uv-vu)x$. Since $x\in U$, $uv-vu\in U$, then $x(uv-vu)+(uv-vu)x\in Z$, thus $2x(uv-vu)\in Z$. Since $x\neq 0$ and since $uv-vu$ has an inverse, $2x(uv-vu)=\lambda\neq 0\in Z$. Multiplying both sides by $uv-vu$, and using that $(uv-vu)^2=-4\lambda_1\lambda_2$, we obtain that $x=\lambda'(uv-vu)$, $\lambda'\in Z$, contradicting that x was linearly independent over Z of u, v and $uv-vu$. Thus no x linearly independent over Z with u, v, $uv-vu$ can be found in U. Thus U *is 3-dimensional over* Z. Also, to be more exact, U has a basis over Z consisting of u, v and $uv-vu$ where $u^2=\lambda_1\neq 0$, $v^2=\lambda_2\neq 0$, $\lambda_1,\lambda_2\in Z$ and where $uv+vu=0$.

Let $N(u)=\{x\in K \mid xu=ux\}$, and we similarly define $N(v)$. By the nature of the basis of U, it is clear that $W=N(u)\cap N(v)$ is such that $W=\{x\in K \mid [x,U]=(0)\}$.

We claim that W is a Lie ideal of K; for $s\in W$, $k\in K$, $t\in U$ implies that

$$(xk-kx)t-t(xk-kx)$$
$$=\{x(kt-tk)-(kt-tk)x\}+\{(xt-tx)k-k(xt-tx)\},$$

since $st - ts = 0$ because $t \in U$, the second $\{\ \}$ is 0; since $kt - tk \in U$, the first $\{\ \}$ is also 0. Thus $xk - kx \in W$. Also $W \supset [K, K]$ for $uv - vu \in [K, K]$ but $uv - vu \notin W$. We should like to show that $W \neq (0)$.

If $s \in S$, u and v as previously chosen in U then we claim that $(us - su)v + v(us - su) \in W$. This can be verified readily by noting that $(us - su)v + v(us - su) = -\{u(vs - sv) + (vs - sv)u\}$ and that $u^2 \in Z$ and $v^2 \in Z$. So, if $W = (0)$ then $(us - su)v + v(us - su) = 0$ for all $s \in S$. But then, since in particular $vs - sv \in S$ when $s \in S$.

$$\{u(vs - sv) - (vs - sv)u\}v + v\{u(vs - sv) - (vs - sv)u\} = 0.$$

Since v commutes with $u(vs - sv)$ and with $(vs - sv)u$, the equation reduces to $2v(u(vs - sv) - (vs - sv)u) = 0$. Since v has an inverse in A we obtain that $u(vs - sv) - (vs - sv)u = 0$ for all $s \in S$. But $u(vs - sv) + (vs - vs)u = 0$ since it is in W. Thus $u(vs - sv) = 0$, and since u has an inverse in A, $vs - sv = 0$ for all $s \in S$. Therefore v is in the center of $\bar{S}$, the subring generated by S. Because A is more than 4-dimensional over Z, by Theorem 9 $\bar{S} = A$. Thus $x \in Z$; that is $v \in Z \cap K = (0)$, forcing $v = 0$, a contradiction. In this way we have proved that $W \neq (0)$.

Since $W \supset [K, K]$ and $W \neq (0)$ and it is a Lie ideal of K, W must also be 3-dimensional over Z. Consider $[W, K]$. Since W is a Lie ideal of K, $[W, K] \subset W$, and is also a Lie ideal of K. It can not be that $[W, K] = (0)$ for then $[W, \bar{K}] = (0)$, and since $\bar{K} = A$, $[W, A] = (0)$ which would imply that $W \subset Z \cap K = (0)$. Thus $[W, K]$ must also be 3-dimensional over Z; since it is contained in the 3-dimensional W, it follows that $[W, K] = W$. Thus $W \subset [K, K]$. Similarly $U \subset [K, K]$. As is seen by examining U, $U \cap W = (0)$. Now $W + U$ is a Lie ideal of K and is contained in $[K, K]$. It is 6-dimensional over Z. If $W + U \neq [K, K]$ then our previous discussion shows that its dimension over Z would be 3, which it is not. Thus $W + U = [K, K]$.

Suppose now that $[K, K] \neq K$. We claim that there exists an $s \in S$ so that $us + su \notin [K, K]$. For, if $us + su$ is in $[K, K]$ for all $s \in S$, since $uk - ku \in [K, K]$, for all $k \in K$, we would have that $ua - a^*u^* \in [K, K]$ for all $a \in A$. As in the proof of Theorem 19 this leads to $[K, K] = K$. Thus, $us_0 + s_0u \notin [K, K]$ for some $s_0 \in S$. Let $x_1 = us_0 + s_0u \in K$, $x_1 \notin [K, K]$. Now $vx_1 - x_1v \in U$, so $vx_1 - x_1v = \alpha v + \beta u + \gamma(uv - vu)$ where $\alpha, \beta, \gamma \in Z$, since u, v and $uv - vu$ form a basis of U over Z. Since $(vx_1 - x_1v)v + v(vx_1 - x_1v) = 0$, we see that $\alpha = 0$. If $\beta = 0$, then $v(x_1 - \gamma u) = (x_1 - \gamma u)v$, and so $x_1 - \gamma u \in N(u) \cap N(v) = W \subset [K, K]$, from which $x_1 \in [K, K]$, a contra-

diction. If, on the other hand, $\beta \neq 0$, $v\left(\frac{x_1 - \gamma u}{\beta}\right) - \left(\frac{x_1 - \gamma u}{\beta}\right)v = u$, that is, there is a $t \in K$, $t \notin [K,K]$ with $ut = tu$ and $u = vt - tv$. Now

$$(uv - vu)t - t(uv - vu)$$
$$= \{u(vt - tv) - (vt - tv)u\} + \{(ut - ut)v - v(ut - tu)\} = 0$$

since $ut - tu = 0$ and since $vt - tv = u$. Thus t commutes with $uv - vu = 2uv$ and with u. So, $utv = tuv = uvt$, thus since u has an inverse in A, $tv = vt$. But then $t \in N(u) \cap N(v) = W \subset [K,K]$, a contradiction. Thus the assumption that $[K,K] \neq K$ leads to a contradiction.

So we assume that $[K,K] = K$. Therefore K is 6-dimensional over Z.

Suppose that $s_1, s_2, \cdots, s_5$ in S are linearly independent over Z. Now $s_iu + us_i \in K$ and commute with u for $i = 1, 2, \cdots, 5$. Thus it is easily seen that $us_i + s_iu = \lambda_{i1}u + \lambda_{i2}w_1 + \lambda_{i3}w_2 + \lambda_{i4}w_3$ where the w_i's are a basis of W and where the $\lambda_{ij} \in Z$ for $i = 1, 2, \cdots, 5$. Therefore, for some $\alpha_i \in Z$, not all 0, $\sum_{i=1}^{5} \alpha_i(s_iu + us_i) = 0$. That is, $(\sum \alpha_i s_i)u + u(\sum \alpha_i s_i) = 0$, and $t = \sum \alpha_i s_i \neq 0$ since the s_i are linearly independent over Z. t is of course in S. Thus, given any 5 linearly independent elements in S we can produce from them an element $t \neq 0 \in S$ so that $ut + tu = 0$.

If the dimension of S over Z is larger than 124, we can find 25 groups of independent elements of S each group consisting of 5 members. From each group we get an element t_i, $i = 1, 2, \cdots, 25$ in S with $ut_i + t_iu = 0$ and where the t_i are linearly independent over Z. We split the t_i's into 5 groups of 5 elements in each. As with u, we obtain 5 elements $p_i \in S$, so that $p_iv + vp_i = 0$ for $i = 1, 2, \cdots, 5$, and where the p_i are linearly independent over Z. Since the p_i's are linear combinations of the t_i's, $p_iu + up_i = 0$ for $i = 1, 2, \cdots, 5$. Thus $p_iuv = uvp_i$, $p_ivu = vup_i$, and so $p_i(uv - vu) = (uv - vu)p_i$. However, as we did with u and v, since there are 5 linearly independent p_i's in Z, we can find an element $q \neq 0 \in S$ which is a linear combination of the p_i's for which $q(uv - vu) + (uv - vu)q = 0$. However, since q is a linear combination of the p_i's and since each p_i commutes with $uv - vu$, we have that $q(uv - vu) = (uv - vu)q$. Therefore $2q(uv - vu) = 0$. Since $2(uv - vu)$ has an inverse in A, $q = 0$ must follow, contradicting that $q \neq 0$. Thus we must assume that the dimension of S over Z is less than 125. But then, since $A = S + K$, the dimension of A over Z is at most 131. By the known results for finite-dimensional simple algebras [7, 8], if A is not the 4×4 matrices over a field Z, U must contain $[K,K]$.

We have finally proved

THEOREM 25. *Let A be a simple ring of characteristic $\neq 2$, and suppose that A is more than 16-dimensional over its center $Z \neq (0)$. Suppose further that $\lambda \in Z$ implies that $\lambda^* = \lambda$. If $U \neq (0)$ is a Lie ideal of K then U must contain $[K,K]$.*

Combining Theorems 23, 24 and 25 we obtain the main theorem of this section, namely

THEOREM 26. *Let A be a simple ring of characteristic $\neq 2$, with an involution and suppose that either $Z = (0)$ or that A is more than 16-dimensional over Z, its center; if K is the set of skew elements of A then every Lie ideal, U, of K must satisfy*

(1) *either* $U \subset Z$ *or* (2) $U \supset [K,K]$.

We combined Theorems 23, 24 and 25 to get the general Theorem 26; however, the sum total of information contained in these three theorems separately exceeds that given in the statement of Theorem 26.

We close the paper with

THEOREM 27. *Let A be as in Theorem 26. Then*

$$[[K,K],[K,K]] = [K,K].$$

Proof. Let $U = [[K,K],[K,K]]$. U is certainly a Lie ideal of K. Thus if $U \neq [K,K]$, then it is strictly contained in $[K,K]$, therefore by Theorem 26, $U \subset Z$. Let $a \in [K,K]$, $k \in K$. Thus

$$b = (ak - ka)a - a(ak - ka) \in U \subset Z.$$

Since aka is in K along with k, if we replace k by aka in the expression for b, and simplify the expression, we have that $aba \in Z$. Since $b \in Z$, $aba = a^2 b \in Z$. If $b \neq 0$ it has an inverse in Z, and so $a^2 \in Z$. If $a^2 \notin Z$, then $b = 0$ for all $k \in K$, so $a(ak - ka) = (ak - ka)a$ for all $k \in K$. That is,

$$(1) \qquad a^2k + ka^2 = 2aka \qquad \text{for all } k \in K.$$

Consider $k = as + sa \in K$ where $s \in S$. For this k (1) yields

$$0 = a^2k + ka^2 - 2aka = a^2(as - sa) - (as - sa)a^2,$$

thus

$$(2) \qquad a^2(as - sa) = (as - sa)a^2 \qquad \text{for all } s \in S.$$

Since $a^2(ak - ka) = (ak - ka)a^2$, we obtain $a^2(ar - ra) = (ar - ra)a^2$ for all $r \in A$. Now, since $a^2r - ra^2 = a(ar - ra) + (ar - ra)a$ we have that $a^2(a^2r - ra^2) = (a^2r - ra^2)a^2$ for all $r \in A$. By the sublemma of Theorem 22,

$a^2 \in Z$. Thus in any case $a^2 \in Z$ for all $a \in [K,K]$. If $a^2 = 0$ for $a \in [K,K]$, then since $a(ak-ka)-(ak-ka)a = -2aka \in U \subset Z$, we obtain $aka \in Z$, and since a has no inverse in A, $aka = 0$ for all $k \in K$. As we have shown several times earlier, this forces $a = 0$.

Consequently $a \neq 0 \in [K,K]$ implies that $a^2 = \lambda \neq 0 \in Z$.

Now $Z \ni a(ak-ka)-(ak-ka)a = 2a(ak-ka)$, whence $a(ak-ka) \in Z$ for all $a \in [K,K]$ and all $k \in K$. Since a has an inverse in A, this implies that $a(ak-ka) = (ak-ka)a$ for all $k \in K$. However, since $a^2 \in Z$, $a(ak-ka) + (ak-ka)a = 0$, and so $2a(ak-ka) = 0$. Since a is regular, this forces $ak-ka = 0$ for all $k \in K$. Thus a is in the center of $\bar{K} = A$. That is, $[K,K] \subset Z$.

Repeating the argument used above on K instead of $[K,K]$ this time, we arrive at $K \subset Z$. But then $[K,K] = (0)$ and so $[[K,K],[K,K]] = [K,K] = (0)$, and the theorem is proved.

UNIVERSITY OF PENNSYLVANIA.

REFERENCES.

[1] Willard E. Baxter, "Lie simplicity of a special class of associative ring," to appear.

[2] J. Dieudonné, "On the structure of unitary groups," *Transactions of the American Mathematical Society*, vol. 72 (1952), pp. 367-385.

[3] I. N. Herstein, "On the Lie and Jordan rings of a simple associative ring," *American Journal of Mathematics*, vol. 77 (1955), pp. 279-285.

[4] ———, "On the Lie ring of a division ring," *Annals of Mathematics*, vol. 60 (1954), pp. 571-575.

[5] ———, "On the Lie ring of a simple ring," *Duke Mathematical Journal*, vol. 22 (1955), pp. 471-476.

[6] N. Jacobson, "Structure theory for algebraic algebras of bounded degree," *Annals of Mathematics*, vol. 46 (1945), pp. 695-707.

[7] ———, "Classes of restricted Lie algebras of characteristic p. I," *American Journal of Mathematics*, vol. 63 (1941), pp. 481-515.

[8] I. Kaplansky, "Report on seminar on simple Lie algebras," *Bulletin of the American Mathematical Society*, vol. 60 (1954), pp. 470-471.

[9] J. Levitski, "A problem of A. Kurosh," *Bulletin of the American Mathematical Society*, vol. 52 (1946), pp. 1033-1035.

[10] ———, "On the radical of a general ring," *ibid.*, vol. 49 (1943), pp. 462-466.

PARTIAL DIFFERENCE SETS.*

By D. R. Hughes.[1]

1. Introduction. The notion of a transitive projective plane and the resulting characterization of the plane by a group with a difference set have been the subject of much interest in recent years (see [5, 9]). In [10] a somewhat similar situation, a transitive affine plane, is studied, and in [11] another similar situation arose in the investigation of associative planar division neo-rings. In this paper we introduce a generalization of the above situations which includes quite a number of other types of projective planes. In each case the plane is characterized by a (collineation) group with a particular subgroup structure and a subset called a "partial difference set."

2. Planar ternary rings. Some use will be made of the planar ternary rings developed by Hall ([8]), but with a different coordinatizing scheme and a different notation; in particular, the ternary function F and the coordinatizing scheme of [11] will be used. In [8] Hall has developed a number of equivalences between the algebraic structure of the planar ternary ring and the geometric structure of the plane, and in [13] more equivalences of this type will be found (indeed, [13] is a very good source for all of these equivalences). All of these results, with perhaps slight modifications to account for the different coordinatizing scheme, carry over to the scheme used here, and those that will be needed will be listed. First a brief sketch of the ternary ring will be given.

The *planar ternary ring* (R, F) is a set R containing at least the two distinct elements 0 (zero) and 1 (one), together with a *ternary function* F (mapping ordered triples of R upon R) satisfying:

(A) $F(a, 0, c) = F(0, b, c) = c$, all $a, b, c \in R$;

(B) $F(a, 1, 0) = F(1, a, 0) = a$, all $a \in R$;

* Received September 14, 1955; revised, April 20, 1956.

[1] This research was carried out while the author was a National Science Foundation Fellow.

(C) if $a, b, c, d \in R$, $a \neq c$, then there is a unique $x \in R$ such that $F(x, a, b) = F(x, c, d)$;

(D) if $a, b, c \in R$, then there is a unique $x \in R$ such that $F(a, b, x) = c$;

(E) if $a, b, c, d \in R$, $a \neq c$, then there is a unique ordered pair $x, y \in R$ such that $F(a, x, y) = b$, $F(c, x, y) = d$.

For all $a, b \in R$, $a \cdot b$ or ab is defined to be $F(a, b, 0)$, and $a + b$ is defined to be $F(1, a, b)$. Then the set R^* of non-zero elements of R is a loop under the operation $(\cdot)$ with identity 1, and R is a loop under the operation $(+)$ with identity 0; these are the multiplicative and additive loops, respectively, of (R, F). If $F(a, b, c) = ab + c$ for all $a, b, c \in R$ then the ring is said to be *linear*.

In the coordinate scheme of [11] the plane π is related to (R, F) as follows: points are (a, b), (a), (∞), for all $a, b \in R$; lines are $[m, k]$, $[\infty, (k, 0)]$, L_x, for all $m, k \in R$. The rules of incidence are: (a, b) is on $[m, k]$ if $F(m, a, b) = k$, and (a, b) is on $[\infty, (a, 0)]$; (m) is on $[m, k]$ and L_x; (∞) is on $[\infty, (k, 0)]$ and L_∞. Note that this scheme is different from the scheme in [13], as well as that in [8].

In any projective plane if P is a point and L is a line, then the plane is said to be (P, L) *transitive* (*with group* H) if there is a group H of collineations which fixes every point on L and every line through P and which is transitive and regular on the "non-fixed" points on any line through P. That is, if Q and R are points collinear with P, both distinct from P and neither on L, then there is a unique $h \in H$ such that $Qh = R$. For further details about (P, L) transitivity, see [2, 13]; we remark that a plane is (P, L) transitive if and only if Desargues' Theorem is valid with P as the center of perspectivity and L as the line of perspectivity.

Now assume that π is coordinatized by the planar ternary ring (R, F) with $X = (0)$, $Y = (\infty)$, $O = (0, 0)$.

THEOREM 1. *(R, F) is linear with associative addition if and only if π is (Y, XY) transitive with group H; furthermore, the additive loop of (R, F) is then isomorphic to H.*

THEOREM 2. *(R, F) is linear with associative multiplication if and only if π is (X, OY) transitive with group H; furthermore, the multiplicative loop of (R, F) is then isomorphic to H.*

We define the *left distributive law* to be $a(b + c) = ab + ac$, and say that (R, F) is *left distributive* if this law holds for all elements of R. The right distributive law is defined analogously, as is right distributivity.

THEOREM 3. *(R,F) is linear, left distributive, and has associative addition if and only if π is (P,XY) transitive for every point P on the line XY.*

THEOREM 4. *(R,F) is linear, left distributive (right distributive), and has associative multiplication if and only if π is (X,OY) and (Y,OX) transitive ((X,OY) and (O,XY) transitive).*

The proofs of the above theorems can be found in [13], or can be obtained by modifying or extending the proofs slightly. A planar ternary ring satisfying the hypotheses of Theorem 3 will be called a *left Veblen-Wedderburn system*, or merely a left V-W system (and a right V-W system is defined similarly), and a left V-W system with associative multiplication will be called a *left near-field.*

3. Partially transitive planes. Throughout the rest of this paper "projective plane" will always mean "finite projective plane." Then if π is a projective plane there will be an integer n (called the *order* of π) such that every point (line) is incident with $n+1$ lines (points) and such that π contains a total of n^2+n+1 points (lines).

Suppose π is a projective plane of order n, and $\mathfrak{G}$ is a non-trivial group of collineations of π; let π_0 be the set of points and lines of π that are fixed (element-wise) by every collineation of $\mathfrak{G}$. Let the points and lines of π_0 be called fixed points and fixed lines; let the points (lines) of π that are not in π_0 but are on lines (contain points) of π_0 be called tangent points (tangent lines); let the remaining points (lines) be called ordinary points (ordinary lines). Finally, suppose that $\mathfrak{G}$ is transitive and regular on both the ordinary points and the ordinary lines; i.e., if X, Y is an ordered pair of ordinary points (or lines) then there is a unique $g \in \mathfrak{G}$ such that $Xg = Y$. Then we shall say that π is a *partially transitive and regular plane with respect to $\mathfrak{G}$ and π_0*, or merely that π is *partially transitive.*

Obviously π_0 is either a degenerate subplane or a (non-degenerate) subplane of π (see [8] for a list of all degenerate planes). Since $\mathfrak{G}$ is transitive and regular on both ordinary points and ordinary lines, the number of ordinary points equals the number of ordinary lines (and this common number is the order of $\mathfrak{G}$); hence it is easy to see that π_0 contains the same number of points and lines. Thus π_0 must be one of the following types:[2]

[2] In the abstract of the paper "Partial difference sets," presented to the American Mathematical Society on Oct. 22, 1955, the author stated essentially that type $(4,m)$ only occurs if $m=3$, and that types $(5,m)$, $(6,m)$ never occur. This is incorrect, and

(0) π_0 is empty.

(1a) π_0 consists of a point Q_0 and a line K_0, Q_0 on K_0.

(1b) π_0 consists of a point Q_0 and a line K_0, Q_0 not on K_0.

(2) π_0 consists of two points Q_0 and Q_1 and two lines K_0 and K_1, where $K_0 = Q_0Q_1$ and Q_0 is on K_1.

(3) π_0 consists of three non-collinear points Q_i, $i = 0, 1, 2$, and the three lines $K_0 = Q_1Q_2$, $K_1 = Q_0Q_2$, $K_2 = Q_0Q_1$.

$(4, m)$ π_0 consists of m points $(m \geqq 3)$ Q_i, $i = 1, 2, \cdots, m$, on a line K_0, a point Q_0 not on K_0, and the $m+1$ lines K_0, $K_i = Q_0Q_i$.

$(5, m)$ π_0 consists of $m+1$ points $(m \geqq 2)$ Q_i, $i = 0, 1, \cdots, m$ on a line K_0, and $m+1$ lines K_i, $i = 0, 1, \cdots, m$, each through Q_0.

$(6, m)$ π_0 is a subplane (i.e., non-degenerate) of order m, with points Q_i, lines K_i, $i = 0, 1, \cdots, m^2 + m$.

Types (1a), (1b), (2), (3) would be special cases of types $(4, m)$ and $(5, m)$ if we allowed $m < 3$ or $m < 2$; however, the analysis is quite different in case $m \geqq 3$ or $m \geqq 2$, respectively, so we separate these cases. The results in this section will apply to all of the types, and we do not distinguish them until subsequent sections.

LEMMA 1. *Every tangent line contains ordinary points.*

Proof. Suppose, if possible, that L is a tangent line containing the fixed point Q and n tangent points $A_1, A_2, \cdots, A_n$. Through each point A_i there is a fixed line K_i and the intersection of two such fixed lines is a fixed point. If all the lines K_i contain the single point B then (since BQ is certainly a fixed line) every line through B is a fixed line, and so there are no ordinary points in the plane. This contradicts the assumption that $\mathfrak{G}$ is non-trivial. Note that if $n = 2$ then there is a point B on all these lines K_i, since there are only two such lines. So we can assume that all the K_i do not pass through a single point, and that $n > 2$. Hence it is easy to see that π_0 contains a set of four points, no three of them collinear, and thus π_0 is a non-degenerate subplane of π, of order m, say. If $n = m^2$, then every point of π is on a line of π_0, and this would contradict the assumption that $\mathfrak{G}$ is non-trivial. Thus (see, for instance, [5]) we know that $n \geqq m^2 + m$. Since π_0 contains at least the n lines K_i, and since Q is on at least 4 lines of π_0,

---

an example of type $(4, m)$ with $m = 4$ is given in Section 5; infinitely many examples of type $(5, m)$ are given in Section 6. The author has no examples of type $(6, m)$ but has no proof that they cannot occur.

π_0 must contain at least $n+4$ lines; so $m^2+m+1 \geqq n+4$. This contradicts $n \geqq m^2+m$, and the lemma is proven.

Now let P_0 be an arbitrarily chosen ordinary point, and let J_0 be an arbitrarily chosen ordinary line. Let D be the set of all $d \in \mathfrak{G}$ such that $P_0 d$ is on J_0. Let $R_i = J_0 K_i$ and $L_i = P_0 Q_i$; let $\mathfrak{R}_i$ be the subgroup of $\mathfrak{G}$ which fixes the point R_i, and let $\mathfrak{L}_i$ be the subgroup of $\mathfrak{G}$ which fixes the line L_i. Since $\mathfrak{R}_i$ consists exactly of those $x \in \mathfrak{G}$ such that $J_0 x$ contains R_i, $\mathfrak{R}_i$ has order equal to the number of ordinary lines through R_i; similarly $\mathfrak{L}_i$ has order equal to the number of ordinary points on L_i. By Lemma 1 each tangent line through Q_i contains an ordinary point and is thus an image $L_i x$ of L_i; so it is fixed by the group $x^{-1}\mathfrak{L}_i x$. Similarly for the tangent points on K_i: each is fixed by a conjugate of $\mathfrak{R}_i$. Also, it is clear that $\mathfrak{R}_i \cap \mathfrak{R}_j = \mathfrak{L}_i \cap \mathfrak{L}_j = 1$ if $i \neq j$; for otherwise there would be a non-identity element of $\mathfrak{G}$ which would fix $P_0 = L_i L_j$, or $J_0 = R_i R_j$.

LEMMA 2. *If Q_i is not on K_j, then $\mathfrak{L}_i$ and $\mathfrak{R}_j$ are conjugates in $\mathfrak{G}$; thus if there is a line of π_0 which does not contain either Q_i or Q_j then $\mathfrak{L}_i$ and $\mathfrak{L}_j$ are conjugates in $\mathfrak{G}$.*

Proof. It is only necessary to prove the first part of the lemma. Suppose Q_i not on K_j; then the line $L = Q_i R_j$ is a tangent line, so $L = L_i x$ for some $x \in \mathfrak{G}$, by Lemma 1. Hence R_j is fixed by both $\mathfrak{R}_j$ and $x^{-1}\mathfrak{L}_i x$, so $\mathfrak{R}_j = x^{-1}\mathfrak{L}_i x$.

THEOREM 5. (a) *If $g \in \mathfrak{G}$, $g \notin \mathfrak{L}_i$ for any i, then $g = d_1 d_2^{-1}$ for a unique ordered pair $d_1, d_2 \in D$; if $g \in \mathfrak{L}_i$ for some i, $g \neq 1$, then $g \neq d_1 d_2^{-1}$ for any $d_1, d_2 \in D$.*

(b) *If $g \in \mathfrak{G}$, $g \notin \mathfrak{R}_i$ for any i, then $g = d_1^{-1} d_2$ for a unique ordered pair $d_1, d_2 \in D$; if $g \in \mathfrak{R}_i$ for some i, $g \neq 1$, then $g \neq d_1^{-1} d_2$ for any $d_1, d_2 \in D$.*

(c) *$\mathfrak{R}_i \cap \mathfrak{R}_j = \mathfrak{L}_i \cap \mathfrak{L}_j = 1$ if $i \neq j$.*

Proof. Let $g \in \mathfrak{G}, g \neq 1$, and consider the line $P_0 \cdot P_0 g = L$. Either $L = J_0 b$ for a unique $b \in \mathfrak{G}$, or $L = L_i$ for a unique i, and in the latter case (and only then) $g \in \mathfrak{L}_i$. If $L = J_0 b$, then $P_0 b^{-1}$, $P_0 g b^{-1}$ are both on J_0, so $b^{-1} = d_2 \in D$, $g b^{-1} = d_1 \in D$, or $g = d_1 d_2^{-1}$. By a reversal of the above it is clear that d_1 and d_2 are unique and the cases of (a) are mutually exclusive. Similarly, (b) is proven using the point $J_0 \cdot J_0 g$, and (c) has already been demonstrated.

LEMMA 3. *If $\mathfrak{G}$ contains an element b of order two, then $b \in \mathfrak{L}_i \cap \mathfrak{R}_j$ for some i and j.*

Proof. Suppose $b \notin \mathfrak{L}_i$ for any i. Then by Theorem 5, $b = d_1 d_2^{-1}$ for a

unique ordered pair $d_1, d_2 \in D$. However, $b = b^{-1} = d_2 d_1^{-1}$, whence $d_1 = d_2$ and $b = 1$, a contradiction. Similarly, $b \notin \mathfrak{R}_i$ for any i is contradictory.

Of the various types given in this section, we shall henceforth exclude type (0), as that is the type given by Bruck in [5] (and including as an important special case the cyclic difference sets of Hall; see [9]). The remaining types will be analyzed in more detail, and we shall consistently use the terminology and notation introduced in this section. Noting Theorem 5, we shall refer to D as a *partial difference set* for the group $\mathfrak{G}$. Of course, a knowledge of the subgroups $\mathfrak{R}_i$ and $\mathfrak{L}_j$ is also needed for a complete description of the situation. Indeed, the existence of a group $\mathfrak{G}$ with a partial difference set D satisfying (a), (b), (c) of Theorem 5, together with certain conditions on the subgroups $\mathfrak{R}_i$ and $\mathfrak{L}_j$, implies the existence of a projective plane of the appropriate type. These other conditions vary somewhat for the different types, and will be given as they arise.

The following table is the result of straightforward counting, which we omit:

Type	Order of $\mathfrak{G}$	Order of $\mathfrak{R}_0$ and $\mathfrak{L}_0$	Order of $\mathfrak{R}_i$ and $\mathfrak{L}_i$, $i \neq 0$	Number of elements in D
(1a)	n^2	n		n
(1b)	$n^2 - 1$	$n - 1$		n
(2)	$n^2 - n$	n	$n - 1$	$n - 1$
(3)	$(n-1)^2$	$n - 1$	$n - 1$	$n - 2$
$(4, m)$	$(n-1)(n-m+1)$	$n - 1$	$n - m + 1$	$n - m$
$(5, m)$	$n(n-m)$	n	$n - m$	$n - m$
$(6, m)$	$(n-m)(n-m^2)$	$n - m^2$	$n - m^2$	$n - m^2 - m$

Note that types (3) and $(6, m)$ are "symmetric" in the sense that the subgroups for $i = 0$ are not particularly distinguished from those for $i \neq 0$.

4. Types (1a), (1b), (2), (3). These types are somewhat simpler than the others, and generally admit of many examples. With the exception of type (1b), there are both Desarguesian and non-Desarguesian examples of each type known. Furthermore, quite a bit of simplification occurs if $\mathfrak{G}$ is abelian or if certain of the subgroups $\mathfrak{R}_i$ or $\mathfrak{L}_i$ are normal, a situation which cannot occur in the other types. We will first give some examples; in each case the example will be given in terms of a particular planar ternary ring of the plane under consideration.

Type (1a). Let (R, F) be a (not necessarily associative) division ring

(with identity). For each ordered pair $a, b \in R$, consider the mapping $\phi = \phi(a, b)$ defined below:

$$\begin{aligned} \phi\colon (x,y) &\to (x+a, y+ax+b) & [m,k] &\to [m-a, k+ma+b-a^2] \\ (m) &\to (m-a) & [\infty,(k,0)] &\to [\infty,(k+a,0)] \\ (\infty) &\to (\infty) & L_\infty &\to L_\infty. \end{aligned}$$

Then the set of all such mappings is a group $\mathfrak{G}$ of collineations with respect to which the plane is of type (1a), with $Q_0 = (\infty)$, $K_0 = L_\infty$. If we let $P_0 = (0,0)$, $J_0 = [0,0]$, then $\mathfrak{R}_0 = \mathfrak{L}_0 = \{\phi(0,b)\}$ is normal in $\mathfrak{G}$, and $D = \{\phi(a,0)\}$.

Type (1b). Examples of this type will be found in [4].

Type (2). Let (R, F) be a linear planar ternary ring with associative addition and multiplication. For each ordered pair $a, b \in R$, $a \neq 0$, consider the mapping $\phi = \phi(a, b)$ defined below:

$$\begin{aligned} \phi\colon (x,y) &\to (ax, y+b) & [m,k] &\to [ma^{-1}, k+b] \\ (m) &\to (ma^{-1}) & [\infty,(k,0)] &\to [\infty,(ak,0)] \\ (\infty) &\to (\infty) & L_\infty &\to L_\infty. \end{aligned}$$

Then the set of all such mappings is a group $\mathfrak{G}$ of collineations with respect to which the plane is of type (2), with $Q_0 = (\infty)$, $Q_1 = (0)$, $K_0 = L_\infty$, $K_1 = [\infty,(0,0)]$. If we let $P_0 = (1,0)$, $J_0 = [1,0]$, then $\mathfrak{R}_0 = \mathfrak{L}_0 = \{\phi(1,b)\}$ is normal in $\mathfrak{G}$, and $\mathfrak{R}_1 = \mathfrak{L}_1 = \{\phi(a,0)\}$ is normal in $\mathfrak{G}$; D consists of all elements $\phi(a,-a)$.

Type (3). Let (R, F) be a linear planar ternary ring with the left distributive law and associative multiplication. For each ordered pair $a, b \in R$, $a \neq 0$, $b \neq 0$, consider the mapping $\phi = \phi(a, b)$ defined below:

$$\begin{aligned} \phi\colon (x,y) &\to (ax, by) & [m,k] &\to [bma^{-1}, bk] \\ (m) &\to (bma^{-1}) & [\infty,(k,0)] &\to [\infty,(ak,0)] \\ (\infty) &\to (\infty) & L_\infty &\to L_\infty. \end{aligned}$$

Then the set of all such mappings is a group $\mathfrak{G}$ of collineations with respect to which the plane is of type (3), with $Q_0 = (\infty)$, $Q_1 = (0)$, $Q_2 = (0,0)$, $K_0 = [0,0]$, $K_1 = [\infty,(0,0)]$, $K_2 = L_\infty$. If we let $P_0 = (1,1)$, $J_0 = [1,1]$, and let $e \in R$ satisfy $e + 1 = 0$, then $\mathfrak{R}_0 = \mathfrak{L}_0 = \{\phi(1,b)\}$, $\mathfrak{R}_1 = \mathfrak{L}_1 = \{\phi(a,1)\}$, $\mathfrak{R}_2 = \{\phi(a,a)\}$, $\mathfrak{L}_2 = \{\phi(a, eae^{-1})\}$. Each of these subgroups except the last two are normal in $\mathfrak{G}$; D consists of all $\phi(a,b)$ such that $a + b = 1$.

By an obvious modification, if "right distributive" is substituted for

"left distributive" in the above example of type (3), then the plane is still of type (3). Since there exist finite non-associative division rings (for type (1a)) and finite near-fields which are not fields (for types (2) and (3)), it is clear that not only does every Desarguesian plane yield an example of the above types but also that there are non-Desarguesian examples. The examples of type (1b) given in [4] are all Desarguesian, and $\mathfrak{G}$ is cyclic, and it is conjectured that all such finite cyclic examples are Desarguesian (see [10]); the author does not know of any finite non-cyclic examples.

From Lemma 2, the groups $\mathfrak{R}_0$ and $\mathfrak{L}_0$ are conjugate in type (1b); $\mathfrak{R}_1$ and $\mathfrak{L}_1$ are conjugate in type (2); for each i, $\mathfrak{R}_i$ and $\mathfrak{L}_i$ are conjugate in type (3). Furthermore, by a proper choice of P_0 and J_0 it is easy to see that certain of the $\mathfrak{R}_i$ and $\mathfrak{L}_i$ can be made to coincide: for $i=0$ in type (1b), $i=1$ in type (2), and any two values of i in type (3).

THEOREM 6. *If either $\mathfrak{R}_i$ or $\mathfrak{L}_i$ is normal in $\mathfrak{G}$, then $\mathfrak{R}_i=\mathfrak{L}_i$.*

Proof. As noted above, this is a corollary of Lemma 2 for every case except $i=0$ in types (1a) and (2), so we restrict attention to these cases. If $\mathfrak{R}_0$ is normal and $x\in\mathfrak{R}_0$, $x\neq 1$, then certainly $x\notin\mathfrak{L}_1$ (in type (2)), since the orders of $\mathfrak{R}_0$ and $\mathfrak{L}_1$ are relatively prime. If also $x\notin\mathfrak{L}_0$ then $x=d_1d_2^{-1}$ for some $d_1, d_2\in D$, whence $d_2^{-1}xd_2=d_2^{-1}d_1\in\mathfrak{R}_0$, since $\mathfrak{R}_0$ is normal. This is impossible, by (b) of Theorem 5, so we must have $x\in\mathfrak{L}_0$, whence $\mathfrak{R}_0=\mathfrak{L}_0$.

THEOREM 7. *In types* (1a) *and* (2), *if $\mathfrak{R}_0$ is normal in $\mathfrak{G}$, then π is coordinatizable by a linear planar ternary ring with associative addition.*

Proof. If $\mathfrak{R}_0$ is normal, then every tangent point R_0x on K_0 is fixed by $\mathfrak{R}_0=x^{-1}\mathfrak{R}_0x$, whence it is immediate that π is (Q_0, K_0) transitive. Thus the theorem follows from Theorem 1.

THEOREM 8. *In type* (1b) *if $\mathfrak{R}_0$ is normal in $\mathfrak{G}$, or in type* (2) *if $\mathfrak{R}_1$ is normal in $\mathfrak{G}$, then π is coordinatizable by a linear planar ternary ring with associative multiplication.*

Proof. Analogous to the proof of Theorem 7, using Theorem 2 instead of Theorem 1.

THEOREM 9. *In type* (2), *if both of the $\mathfrak{R}_i$ are normal in $\mathfrak{G}$ then π is coordinatizable by a linear planar ternary ring with associative addition and multiplication.*

Proof. Immediate from Theorems 7 and 8, noting that the same ternary ring can be used in the conclusion of each of these theorems.

Thus, referring to the example of type (2) given at the beginning of this section, Theorem 9 actually affords us a necessary and sufficient condition that a plane be coordinatizable by a linear ring with both operations associative.

THEOREM 10. *In type* (3), *if* $\mathfrak{R}_0$ *is normal in* $\mathfrak{G}$ *then* π *is coordinatizable by a linear planar ternary ring with associative multiplication. If* $\mathfrak{R}_0$ *and* $\mathfrak{R}_1$ *are both normal in* $\mathfrak{G}$ *then* π *is coordinatizable by a linear planar ternary ring with the left distributive law and associative multiplication; furthermore,* $\mathfrak{G}$ *is isomorphic to the direct product of* $\mathfrak{R}_0$ *with itself, and all three of the* $\mathfrak{R}_i$ *are isomorphic to one another. If all three of the* $\mathfrak{R}_i$ *are normal in* $\mathfrak{G}$ *then* π *is coordinatizable by a linear planar ternary ring with both distributive laws and associative, commutative multiplication* (*i.e., an abelian planar division neo-ring; see* [11]).

Proof. The first sentence and the first part of the second sentence in the theorem are immediate from Theorems 2 and 4. Assume that $\mathfrak{R}_0$ and $\mathfrak{R}_1$ are normal in $\mathfrak{G}$; it is clear that $\mathfrak{G}$ is their direct product. If $x \in \mathfrak{R}_2$, then $x = ab$, $a \in \mathfrak{R}_0$, $b \in \mathfrak{R}_1$, and this representation is unique; furthermore, for any $a \in \mathfrak{R}_0$, there is exactly one $b \in \mathfrak{R}_1$ such that $ab \in \mathfrak{R}_2$, since $\mathfrak{R}_0 \cap \mathfrak{R}_2 = \mathfrak{R}_1 \cap \mathfrak{R}_2 = 1$. Thus every element of $\mathfrak{R}_2$ can be written in the form $a(aT)$ for some $a \in \mathfrak{R}_0$, where T is a one-to-one mapping of $\mathfrak{R}_0$ upon $\mathfrak{R}_1$. If $a, b \in \mathfrak{R}_0$, then $[a(aT)][b(bT)] = ab[(aT)(bT)] \in \mathfrak{R}_2$, so $(aT)(bT) = (ab)T$. Hence $\mathfrak{R}_0$ is isomorphic to $\mathfrak{R}_1$, and obviously $\mathfrak{R}_2 = \{a(aT)\}$ is also isomorphic to $\mathfrak{R}_0$.

If also $\mathfrak{R}_2$ is normal, let $a, b \in \mathfrak{R}_0$. Then $a^{-1}[b(bT)]a = a^{-1}ba(bT) \in \mathfrak{R}_2$, so $bT = (a^{-1}ba)T$ or $b = a^{-1}ba$. Thus $\mathfrak{R}_0$ is commutative (and so is $\mathfrak{G}$). Since, from Theorem 2, $\mathfrak{R}_0$ is isomorphic to the multiplicative group of the planar ternary ring under consideration, we are done (in view of Theorem 4 and the first part of this theorem).

From Theorem 10 and the example of type (3) given at the beginning of this section, we have a necessary and sufficient condition that π possess a linear coordinate ring with the left distributive law and associative multiplication: π must be of type (3) with two of the $\mathfrak{R}_i$ normal in $\mathfrak{G}$. Following the remark at the end of the example, this is the same as the condition that it possess a linear ring with associative multiplication and the right distributive law (note however, that these are different rings: the points $(0,0)$ and (∞) are interchanged). In fact, if (R, F) is a planar ternary ring which is linear, left distributive, and has associative multiplication, then the interchange of the roles of the points $Y = (\infty)$ and $O = (0,0)$ gives rise to a new ring with the same properties, except that "right distributive"

replaces "left distributive." Thus suppose π is a plane coordinatized by a left near-field (which is not a field); performing the above interchange we arrive at a new system, which however cannot have associative addition. For if it did it would be a right near-field, and it is elementary to see that there would be a collineation moving the line at infinity (L_∞) of the original left near-field coordinate system. It is well-known (see, for instance, [8]) that this implies that the original left near-field is a field, which is contradictory.

The above observations are of interest mainly because they indicate the existence of finite non-trivial examples of what might be called "right (or left) planar division neo-rings": linear planar ternary rings with the right (or left) distributive law, whose addition is not necessarily associative. (In this connection, see [6, 11].) Although the examples given here do not lead to new projective planes, others might, and the lack of associative addition implies that we are not immediately restricted to cases where the order is a prime-power.

We return to the general case of the section now. If ϕ is an automorphism of the group $\mathfrak{G}$ and if $D\phi = Db$ for some $b \in \mathfrak{G}$, then, following Hall ([9]) we call ϕ a *multiplier* of the partial difference set (according to [5], ϕ would be a right multiplier). The concept of multipliers has been a powerful one in the treatment of cyclic difference sets (i.e., cyclic groups of type (0)) and has already been applied to partial difference sets by Hoffman ([10]) and the author ([11]) for types (1b) and (3), respectively; using the techniques of [11], Hoffman's results can be extended from the cyclic case to the abelian case, by the way. Specifically, all of these results are of the following nature: if $\mathfrak{G}$ is abelian, of type (1b) or (3) (or even of type (0)), and if p is any prime divisor of n, then the mapping $\phi: x \to x^p$ is a multiplier. Using this result, abelian examples of type (1b) or (3) appear very likely to be of prime-power order (i.e., n is a power of a prime). However, the same mapping is not even one-to-one for abelian examples of type (1a) or (2), and other difficulties present themselves: in the known proofs for the existence of multipliers, a key step is to show that the element $\Delta = \sum d^{-1}$, for all $d \in D$, is a non-singular element of the group algebra of $\mathfrak{G}$ over the rationals. For types (1a), and (2), Δ is definitely singular, and it does not seem unlikely that this should indicate some kind of fundamentally different situation.

Now we investigate the conditions under which we can construct the plane from the group $\mathfrak{G}$. In each of the types all of the left cosets $d\mathfrak{R}_i$, $d \in D$, are distinct. For if $d_1 = d_2 r$, $r \in \mathfrak{R}_i$, $d_1, d_2 \in D$, then $r = d_2^{-1} d_1$ and by Theorem 5, this implies $d_1 = d_2$. Thus, referring to the table at the end of

Section 3, we see that every left coset of $\mathfrak{R}_i$ is of the form $d\mathfrak{R}_i$, $d \in D$, for $i = 0$ in types (1a) and (2), and that all but one left coset of $\mathfrak{R}_i$ is of this form for the remaining values of i. In each of these latter cases, let $q_i \in \mathfrak{G}$ be so chosen that $q_i\mathfrak{R}_i$ is the unique coset not of the form $d\mathfrak{R}_i$; although q_i is not unique, it is determined up to a right multiple by an element of $\mathfrak{R}_i$.

THEOREM 11. *Whenever q_i is defined, $q_i\mathfrak{R}_i = \mathfrak{L}_i q_i$.*

Proof. Choose $b \in \mathfrak{G}$ such that R_i is on $L_i b$. Then $b^{-1}\mathfrak{L}_i b = \mathfrak{R}_i$, or $\mathfrak{L}_i b = b\mathfrak{R}_i$. But $R_i b^{-1}$ is on L_i, so the collection of lines $J_0\mathfrak{R}_i b^{-1}$ all pass through $R_i b^{-1}$; they are all ordinary lines, so none of them contains Q_i, and hence none of them contains P_0. Thus $1 \notin D\mathfrak{R}_i b^{-1}$, or $b \notin D\mathfrak{R}_i$; so $b\mathfrak{R}_i$ is the (unique) left coset of $\mathfrak{R}_i$ which is not of the form $d\mathfrak{R}_i$, $d \in D$. Thus $b\mathfrak{R}_i = q_i\mathfrak{R}_i$, and $q_i = br$, $r \in \mathfrak{R}_i$. So $\mathfrak{L}_i q_i = \mathfrak{L}_i br = b\mathfrak{R}_i r = b\mathfrak{R}_i = q_i\mathfrak{R}_i$.

Indeed, Theorem 11 can be extended to show that $\mathfrak{L}_i q_i$ is the unique right coset of $\mathfrak{L}_i$ which is not representable as $\mathfrak{L}_i d$, $d \in D$. However, we do not need this in what follows.

THEOREM 12. *Besides* (a), (b), (c) *of Theorem* 5, *the following is also satisfied:*

(d) $q_i\mathfrak{R}_i = \mathfrak{L}_i q_i$, *where $q_i\mathfrak{R}_i$ is defined as the unique left coset of $\mathfrak{R}_i$ not representable as $d\mathfrak{R}_i$, $d \in D$, if such a one exists; furthermore, all of the left cosets $d\mathfrak{R}_i$, $d \in D$, are distinct.*

Now if $\mathfrak{G}$ is a group with a subet D and subgroups $\mathfrak{R}_i$ and $\mathfrak{L}_i$, satisfying the numerical conditions of the table at the end of Section 3 for one of the types (1a), (1b), (2), or (3), and if furthermore (a), (b), (c), (d) of Theorems 5 and 12 are satisfied, then we shall say that $(\mathfrak{G}, \mathfrak{R}_i, \mathfrak{L}_i, D)$ is a *partial difference system* (of the appropriate type). We will now show that the existence of a partial difference system implies the existence of a projective plane of the appropriate type. Actually, we will define the plane for each type but only demonstrate that it is a projective plane for type (2), since this is fairly typical of each of the types; the demonstration for the other types is straightforward.

We define a set π of points and lines, with an incidence relation as below.

Points: (a), for each $a \in \mathfrak{G}$; $(\mathfrak{R}_i a)$ for each right coset of $\mathfrak{R}_i$; Q_i, where i runs over the appropriate integers.

Lines: $[Db]$, for each $b \in \mathfrak{G}$; $[\mathfrak{L}_j b]$, for each right coset of $\mathfrak{L}_i$; K_j, where j runs over the appropriate integers.

Incidence:

(a) on $[Db]$ if $a \in Db$; (a) on $[\mathfrak{L}_j b]$ if $a \in \mathfrak{L}_j b$; (a) never on K_j.

$(\mathfrak{R}_i a)$ on $[Db]$ if $b \in \mathfrak{R}_i a$; $(\mathfrak{R}_i a)$ on K_j if $i = j$; $(\mathfrak{R}_i a)$ on $[\mathfrak{L}_j b]$ as below:

Type (1a): never.

Type (1b): if $q_0 a \in \mathfrak{L}_0 b$.

Type (2): never, if $i = 0$ or $j = 0$; for $i = j = 1$, then if $q_1 a \in \mathfrak{L}_1 b$.

Type (3): never, if $i \neq j$; for $i = j$, then if $q_i a \in \mathfrak{L}_i b$.

Q_i never on $[Db]$; Q_i on $[\mathfrak{L}_j b]$ if $i = j$; Q_i on K_j as demanded by π_0.

Now we consider type (2) and show that π, as defined above, forms a projective plane.

Let (a) and (b) be distinct points (i.e., $a \neq b$). Then either $ab^{-1} \in \mathfrak{L}_j$ for some j, or $ab^{-1} = d_1 d_2^{-1}$ for a unique pair $d_1, d_2 \in D$. In the first case we have $a, b \in \mathfrak{L}_j b$, so both (a) and (b) are on $[\mathfrak{L}_j b]$. In the second case we have $d_1^{-1} a = d_2^{-1} b = c$, whence $a, b \in Dc$ and both points are on $[Dc]$. The arguments are easily reversed to show uniqueness.

Consider the points (a) and $(\mathfrak{R}_i b)$. If $i = 0$, then ab^{-1} is in a coset $d\mathfrak{R}_0$, where $d \in D$, so $ab^{-1} = dr$, where $r \in \mathfrak{R}_0$. Let $c = rb$; then $a \in Dc$ and $c \in \mathfrak{R}_0$, so (a) and $(\mathfrak{R}_0 b)$ are on $[Dc]$. If $i = 1$ then either $ab^{-1} \in d\mathfrak{R}_1$ for some $d \in D$, or $ab^{-1} \in q_1 \mathfrak{R}_1$. The first case is handled exactly the same as when $i = 0$. In the second case we have $ab^{-1} \in q_1 \mathfrak{R}_1 = \mathfrak{L}_1 q_1$, and we let $c = q_1 b$. Then $a \in \mathfrak{L}_1 c$ and $q_1 b \in \mathfrak{L}_1 c$, so both points are on the line $[\mathfrak{L}_1 c]$. Again the arguments are readily reversed to demonstrate uniqueness.

Clearly the points (a) and Q_i are on $[\mathfrak{L}_i a]$, and on no other.

Let $(\mathfrak{R}_i a)$ and $(\mathfrak{R}_j b)$ be distinct points (i.e., $\mathfrak{R}_i a \neq \mathfrak{R}_j b$). Then if $i = j$ both points are on the line K_i. Let $i \neq j$, and note that $\mathfrak{R}_j \mathfrak{R}_i = \mathfrak{G}$, since the two subgroups intersect in the identity; for the same reason, every element of $\mathfrak{G}$ has a unique representation in the form $r_j^{-1} r_i$, where $r_j \in \mathfrak{R}_j$, $r_i \in \mathfrak{R}_i$. Let $ba^{-1} = r_j^{-1} r_i$; then $r_i a = r_j b = c$, whence $c \in \mathfrak{R}_i a$, $c \in \mathfrak{R}_j b$, and so $(\mathfrak{R}_i a)$ and $(\mathfrak{R}_j b)$ are both on $[Dc]$. Again, the uniqueness is straightforward.

Consider the points $(\mathfrak{R}_i a)$ and Q_j. If $i \neq j$, then both points are on K_i, while if $i = j = 0$, both points are on K_0. If $i = j = 1$, then let $b = q_1 a$; we have $q_1 a \in \mathfrak{L}_1 b$, so $(\mathfrak{R}_1 a)$ is on $[\mathfrak{L}_1 b]$ and certainly Q_1 is on $[\mathfrak{L}_1 b]$. Uniqueness is obvious.

Finally, the points Q_0 and Q_1 are both on K_0.

Since π is finite and every line clearly contains $n + 1$ points, this is sufficient to prove that π is a projective plane (the nondegeneracy is trivial if $n > 1$). In order to show that π is of type (2), consider mappings of the

form $(a) \to (ax)$, $(\mathfrak{R}_i a) \to (\mathfrak{R}_i ax)$, etc., for each $x \in \mathfrak{G}$. The set of such mappings forms a group of collineations (isomorphic to $\mathfrak{G}$) with respect to which π is of type (2).

Another kind of example of a partial difference system can be constructed as follows. Let R be a finite field of order $n = 2^t$, and let p be any integer such that both p and $p-1$ are prime to $n-1$ (e.g., $p=2$). Let $\mathfrak{G}$ be the group whose elements are the ordered pairs (a, b), $a, b \in R$, $a \neq 0$, with composition $(a, b)(c, d) = (ac, bc + d)$ (then in fact, $\mathfrak{G}$ is isomorphic to the holomorph of the additive group of R with the automorphism group of multiplications). The identity of G is $(1, 0)$, and $(a, b)^{-1} = (a^{-1}, ba^{-1})$. Let D be the subset of $\mathfrak{G}$ consisting of all elements (x, x^p), $x \neq 0$. Then, as both p-th and $(p-1)$-th roots exist and are unique in R, it is easy to see that every element of $\mathfrak{G}$, excepting the elements in the subgroups $\{(1, b)\} = \mathfrak{R}_0 = \mathfrak{L}_0$, and $\{(a, 0)\} = \mathfrak{R}_1 = \mathfrak{L}_1$, can be represented uniquely as $\delta_1 \delta_2^{-1}$ and as $\delta_3^{-1} \delta_4$, where the δ_i are in D. Further, (c) and (d) of Theorems 5 and 12 are trivially satisfied.

Thus we have a plane π of type (2), of order $n = 2^t$, with nonabelian $\mathfrak{G}$. Since $\mathfrak{R}_0$ is normal, π possesses a coordinate ring which is linear and has associative addition. However, for $n = 2, 4, 8$, π must be Desarguesian (since all planes of order $\leqq 8$ are known to be Desarguesian), although the author does not know if this is generally true. Also, this gives an example of type (2) in which not all of the subgroups $\mathfrak{R}_i$ are normal (since $\mathfrak{R}_1$ is certainly not).

5. Type (4, *m*). Throughout this section we asume that π is of type $(4, m)$, where $m \geqq 3$. Since J_0 can be chosen so that R_0, P_0, Q_0 are collinear, we can assume that $\mathfrak{R}_0 = \mathfrak{L}_0$ and (unless stated to the contrary) we shall always make this assumption. From Lemma 2, all the $\mathfrak{R}_i$ and $\mathfrak{L}_i$, for $i \neq 0$, are conjugates, and hence $\mathfrak{G}$ is certainly not abelian, nor are any of the $\mathfrak{R}_i$ or $\mathfrak{L}_i$, $i \neq 0$, even normal in $\mathfrak{G}$: for $\mathfrak{R}_i \cap \mathfrak{R}_j = \mathfrak{L}_i \cap \mathfrak{L}_j = 1$ if $i \neq j$, and $\mathfrak{R}_i$ or $\mathfrak{L}_j$ do not have order one (see the table in Section 3). Let $q = (n-1)/(m-2)$; we shall show that q is an integer.

THEOREM 13. *If $i \neq 0$ and $a \in \mathfrak{G}$, then $a^{-1}\mathfrak{R}_i a$ fixes a subplane π_1 of π, of order $m-1$; $a^{-1}\mathfrak{R}_i a$ fixes exactly $m-2$ tangent points on any line K_j, $j \neq 0$, and q is an integer. $\mathfrak{R}_i$ has exactly q distinct conjugates, any two of which intersect in the identity, and any two of which fix different subplanes of order $m-1$.*

Proof. Let π_1 be the set of points and lines that are fixed by every

element of $a^{-1}\mathfrak{R}_i a$. π_1 contains at least one tangent point on K_i, and so if $j \neq 0$, $a^{-1}\mathfrak{R}_i a$ fixes at least one tangent point on K_j, because there is at least one point Q_k on K_0 which is not on K_i or K_j. Thus π_1 contains at least four points, no three of them collinear, since π_1 certainly contains the m points Q_j, $j \neq 0$; so π_1 is a non-degenerate subplane of π. But π_1 contains only the m points Q_j, $j \neq 0$, from among the points on K_0, and so π_1 must contain m points on all of its lines: thus its order is $m-1$.[3] Besides Q_0 and Q_j, $j \neq 0$, π_1 contains then exactly $m-2$ tangent points on the line K_j. If $j = i$, then these $m-2$ points are just the points $R_i a x$, where x is in the normalizer of $a^{-1}\mathfrak{R}_i a$ in $\mathfrak{G}$. If $a^{-1}\mathfrak{R}_i a \neq b^{-1}\mathfrak{R}_i b$, but there is a non-identity element in common to these conjugate subgroups, then this element must fix exactly m points on K_0, but more than m points on K_i, since it fixes the points that are fixed by $a^{-1}\mathfrak{R}_i a$ and also the points fixed by $b^{-1}\mathfrak{R}_i b$, and these two sets of points (on K_i) cannot be the same (for otherwise the two subgroups would be equal).

Thus distinct conjugates of $\mathfrak{R}_i$ intersect in the identity, and in fact, fix sets of tangent points on K_i which are distinct. So the $n-1$ tangent points on K_i break up into sets of $m-2$ points each, hence q is an integer. In fact, since there are q sets of such tangent points on K_i and each corresponds to a different conjugate of $\mathfrak{R}_i$, q must be the index of the normalizer of $\mathfrak{R}_i$ in $\mathfrak{G}$.

If $\mathfrak{A}$ is any subgroup of $\mathfrak{G}$, let $N(\mathfrak{A})$ be the normalizer of $\mathfrak{A}$ in $\mathfrak{G}$. Using the new parameter q in place of n, $\mathfrak{G}$ has order $q(q-1)(m-2)^2$, $\mathfrak{R}_0$ (and $\mathfrak{L}_0$) has order $q(m-2)$, $\mathfrak{R}_i$ and $\mathfrak{L}_i$, $i \neq 0$, have order $(q-1)(m-2)$, $N(\mathfrak{R}_i)$ and $N(\mathfrak{L}_i)$ have order $(q-1)(m-2)^2$. Since π, of order n, possesses a subplane of order $m-1$, we must have $n = (m-1)^2$ or $n \geqq (m-1)^2 + (m-1) = m^2 - m$. In the first case we have $q = m$ and in the second case $q \geqq (m^2 - m - 1)/(m-2) = m + 1 + 1/(m-2)$, so if $q \neq m$, then $q \geqq m + 2$.

THEOREM 14. *Both m and n are odd, unless $n = (m-1)^2$.*

Proof. If m is even, then $n - 1 \equiv m - 2 \equiv 0 \pmod 2$, so n is odd and $n - m + 1$ is even. So $\mathfrak{R}_i$, $i \neq 0$, possesses an element of order two, and so does each of its conjugates. Thus by Lemma 3, every conjugate of $\mathfrak{R}_i$ must be an $\mathfrak{R}_j$, where $j \neq 0$, and so $q = m$ and $n = (m-1)^2$.

If m is odd, then $n - m + 1 \equiv n \pmod 2$ so either $\mathfrak{R}_0$ or $\mathfrak{R}_i$, $i \neq 0$,

[3] For if $a^{-1}\mathfrak{R}_i a$ fixed a point S on K_0, $S \neq Q_i$ for any i, then since $a^{-1}\mathfrak{R}_i a$ fixes a tangent point T on K_i, $a^{-1}\mathfrak{R}_i a$ would fix the line $L = ST$; but L is an ordinary line, and this is contradictory.

has even order and possesses an element of order two. If n is even then it must be $\mathfrak{R}_i$, $i \neq 0$, that possesses this element, so as above, $q = m$ and $n = (m-1)^2$.

Now let $\mathfrak{H} = N(\mathfrak{R}_1) \cap \mathfrak{R}_0$.

THEOREM 15. *The order of $\mathfrak{H}$ is $m-2$.*

Proof. Consider the $m-2$ points $R_1 a$, where $a \in N(\mathfrak{R}_1)$. Let $i \neq 0, 1$, and let S be the intersection of the lines L_0 and $R_1 Q_i$. Each of the points $L_0(R_1 a \cdot Q_i)$, for $a \in N(\mathfrak{R}_1)$, is a point Sb, where $b \in \mathfrak{R}_0$; but clearly, for each such b, we also have $b \in N(\mathfrak{R}_1)$, so $b \in \mathfrak{H}$. Conversely, any element in $\mathfrak{H}$ must be one of the elements b defined in this way, so $\mathfrak{H}$ has order $m-2$.

THEOREM 16. *If $n \neq (m-1)^2$, then both $\mathfrak{H}$ and $\mathfrak{R}_0$ are normal in $\mathfrak{G}$. Furthermore, every element of $\mathfrak{R}_0$ which is not in $\mathfrak{H}$ has order two and $q = 2^t$ for some t.*

Proof. Suppose $b \in \mathfrak{R}_0$, $b \notin \mathfrak{H}$, and $r \in \mathfrak{R}_1$, $r \neq 1$; suppose also that $b^{-1} r b = r$. Then r fixes the point $R_1 b$, whereas (since $b \notin N(\mathfrak{R}_1)$) $R_1 b$ is not one of the $m-2$ tangent points on K_1 that are fixed by $\mathfrak{R}_1$. This is impossible (see the proof of Theorem 13). So if $x, y \in \mathfrak{R}_1$ and $x^{-1} b x = y^{-1} b y$, we have $b^{-1}(y x^{-1}) b = y x^{-1}$, and thus $x = y$. Hence all the $(q-1)(m-2)$ conjugates of b by elements of $\mathfrak{R}_1$ are distinct. Now if $n \neq (m-1)^2$ then $\mathfrak{R}_0$ has even order $n-1$ while $\mathfrak{H}$ has odd order $m-2$, by Theorem 14; so $\mathfrak{R}_0$ possesses an element b of order two, and $b \notin \mathfrak{H}$. The $(q-1)(m-2)$ conjugates of b by elements of $\mathfrak{R}_1$ are all distinct and they all have order two, so by Lemma 3 they must all be in $\mathfrak{R}_0$ (for $\mathfrak{R}_i$, $i \neq 0$, has odd order). So all the elements of $\mathfrak{R}_0$ not in $\mathfrak{H}$ have order two, and there are no other elements of order two in $\mathfrak{G}$. Then any two conjugates of $\mathfrak{R}_0$ must intersect in a group containing these $(q-1)(m-2)$ elements of order two; it is easy to see that this set of elements generate $\mathfrak{R}_0$, so $\mathfrak{R}_0$ is normal in $\mathfrak{G}$. The group $\mathfrak{H}$ consists exactly of all the elements of $\mathfrak{R}_0$ which have odd order (plus the identity), so $\mathfrak{H}$ is normal in $\mathfrak{R}_0$, and even in $\mathfrak{G}$, since $\mathfrak{H}$ is in fact an invariant subgroup of $\mathfrak{R}_0$. Finally, $\mathfrak{R}_0/\mathfrak{H}$ has order q and must also be elementary abelian of exponent two; this finishes the proof.

Now we prove a lemma about arbitrary groups which enables us to classify further our partial difference system.

LEMMA 4. *If G is a (finite or infinite) group, and if H is the subgroup generated by all of the elements which do not have order two, then either $H = 1$, $H = G$, or H has index two in G.*

Proof. Suppose $H \neq 1$, $H \neq G$. If $a \in G$, $a \notin H$, then a has order two; if furthermore, $b \in G$, $b \notin H$, then if $ab \notin H$, $abab = 1$, so $ab = b^{-1}a^{-1} = ba$. If $h \in H$, then since $ah \notin H$, $(ah)^2 = 1$, so $aha = h^{-1}$. Now suppose $a, b \in G$, $a, b, ab \notin H$. For any $h \in H$,

$$h^{-1} = (ab)h(ab) = (ba)h(ab) = b(aha)b = bh^{-1}b = h;$$

i.e., for any $h \in H$, $h^2 = 1$. Since H certainly contains elements (different from the identity) whose order is not two, this is contradictory, whence if $a, b \in G$, $a, b \notin H$, then $ab \in H$. Thus H has index two in G.

THEOREM 17. *In type* $(4, m)$, *if* $m \neq 3$, *then* $n = (m-1)^2$.

Proof. From Theorem 16, $\mathfrak{H}$ consists of all the elements of $\mathfrak{R}_0$ whose order is not two (plus the identity), if $n \neq (m-1)^2$. So, by Lemma 4 either $\mathfrak{H} = 1$, which means $m = 3$, or $\mathfrak{H}$ has index two in $\mathfrak{R}_0$, since $\mathfrak{H}$ is certainly not equal to $\mathfrak{R}_0$. So if $m \neq 3$, then q, which is the index of $\mathfrak{H}$ in $\mathfrak{R}_0$, is two, and this is impossible, since $q \geqq m$.

Thus we have only two cases left. If $n \neq (m-1)^2$, then $m = 3$ and $n = 2^t + 1$, and a good deal more can be said about the group $\mathfrak{G}$ and the plane π. For $\mathfrak{R}_0$ is an elementary abelian group of order 2^t and is normal in $\mathfrak{G}$; for $i \neq 0$, $\mathfrak{R}_i$ defines a transitive and regular automorphism group of $\mathfrak{R}_0$. Thus $\mathfrak{R}_0$ is isomorphic to the additive group of a right near-field (which might be a field). $\mathfrak{R}_1$, say, is isomorphic to the multiplicative group of the same near-field, and $\mathfrak{G}$ is isomorhpic to the holomorph of the additive group of the near-field with the automorphism group of right multiplications. Furthermore, since π contains subplanes of order $m-1=2$, and since π has odd order, π is never Desarguesian. We shall return to this case later to give a more complete description of $\mathfrak{G}$ as the holomorph mentioned above.[4]

Now we shall investigate the converse problem; i.e., find the "axioms" corresponding to Theorem 12, for type $(4, m)$. It is apparent that part of (d) still holds: all of the cosets $d\mathfrak{R}_i$, $d \in D$, are distinct. Since we have chosen R_0 on L_0, it is easy to prove that the unique left coset of $\mathfrak{R}_0$ which is not of the form $d\mathfrak{R}_0$, $d \in D$, is $\mathfrak{R}_0$ itself. In what follows, it is assumed

[4] Let H be a finite group, written additively, of order k, and H' a group of automorphisms of H, transitive and regular on the non-zero elements of H (whence H' has order $k-1$). Let $e \neq 0$ be any fixed element of H, and define a multiplication in H as follows: (i) $x0 = 0$, all $x \in H$; (ii) if $b \neq 0$, then $xb = x\phi_b$, all $x \in H$, where ϕ_b is the unique element of H' satisfying $e\phi_b = b$. Then it is easy to prove that H, under these two operations, is a right near-field. If G is a group of order $k(k-1)$ containing H and H', such that each automorphism of H by an element of H' is an inner automorphism in G, then G is necessarily the holomorph of H with H'.

that i, j, k are all non-zero. If $i \neq j$, choose $g_{ij} \in \mathfrak{G}$ such that the point R_j is on $L_i g_{ij}$; then $\mathfrak{R}_j = g_{ij}^{-1} \mathfrak{L}_i g_{ij}$. Then since P_0 is on L_j, $P_0 g_{ji}$ is on $L_j g_{ji}$, and since R_i is also on $L_j g_{ji}$, $P_0 g_{ji}$ is not on J_0; for if it were, it would have to be the point R_i. Thus $P_0 g_{ji} r$, for any $r \in \mathfrak{R}_i$, is on $L_j g_{ji} r = L_j g_{ji}$ and is not on J_0. So $P_0 g_{ji} r \neq P_0 d$ for any $d \in D$ and any $r \in \mathfrak{R}_i$, and hence $g_{ji} \mathfrak{R}_i \neq d\mathfrak{R}_i$ for any $d \in D$. By a similar argument, it is easy to demonstrate that all of the left cosets $g_{ji}\mathfrak{R}_i$ (as j varies) are distinct. Then, by counting, it is evident that each left coset of $\mathfrak{R}_i$ which is not of the form $d\mathfrak{R}_i$, $d \in D$, is of the form $g_{ji}\mathfrak{R}_i$ for a unique j. Thus we have the result corresponding to (d) of Theorem 12. But we need more here.

Consider a pair of points $R_i b$ and R_j, $i \neq j$, and let L be the line joining them. Then either L is a line $L_k g_{kj}$ for a unique k (i.e., Q_k, $R_i b$, R_j are collinear), or L is a line $J_0 x$. In the first case, R_i is on $L_k g_{ki}$, so $R_i b$ is on $L_k g_{ki} b = L_k g_{kj}$; hence $g_{kj} \in \mathfrak{L}_k g_{ki} b$, or $g_{kj}\mathfrak{R}_j = g_{ki}\mathfrak{R}_i b$, where k is unique. In the second case, since $R_i b$ is on $J_0 x$, we have $R_i b = R_i x$, and similarly, $R_j x = R_j$. So $x \in \mathfrak{R}_i b \cap \mathfrak{R}_j$, and clearly x is the only element in this intersection.

THEOREM 18. *Besides* (a), (b), (c) *of Theorem* 5, *the following are satisfied:*

(e) *There are elements* $g_{ij} \in \mathfrak{G}$ *for* $i, j \neq 0$, $i \neq j$, *such that every left coset of* $\mathfrak{R}_i$, $i \neq 0$, *can be represented uniquely as* $d\mathfrak{R}_i$, $d \in D$, *or as* $g_{ji}\mathfrak{R}_i$, *and also having the property* $g_{ji}\mathfrak{R}_i = \mathfrak{L}_j g_{ji}$. *Every left coset of* $\mathfrak{R}_0$ *can be expressed uniquely as* $d\mathfrak{R}$, $d \in D$, *excepting the coset* $\mathfrak{R}_0$ *itself.*

(f) *If* $a \in G$, *then* $\mathfrak{R}_i a \cap \mathfrak{R}_0$, $i \neq 0$, *contains a single element. If* $i, j \neq 0$, $i \neq j$, *then either* $\mathfrak{R}_i a \cap \mathfrak{R}_j$ *contains a single element or* $g_{ki}\mathfrak{R}_i a = g_{kj}\mathfrak{R}_j$ *for a unique* k, *but not both.*

Proof. Everything is proven, excepting the first sentence of (f). But this is trivial, since different element of $\mathfrak{R}_0$ cannot be in the same left coset of $\mathfrak{R}_i$.

We are now in a position to define the projective plane from a partial difference system of type $(4, m)$ (i.e., a system $(\mathfrak{G}, \mathfrak{R}_i, \mathfrak{L}_i, D)$ satisfying the numerical conditions of the table at the end of Section 3 and the conditions (a), (b), (c), (e), (f) of Theorems 5 and 18). It is worth noting that all of these conditions are not independent; certainly (b) is not needed, for instance. The plane is defined to consist of points and lines exactly as in Section 4, and incidence is also exactly as in Section 4, excepting for the case $(\mathfrak{R}_i a)$ on $[\mathfrak{L}_j b]$, which is as follows:

(i) if $i = j = 0$, then if $\mathfrak{R}_0 a = \mathfrak{L}_0 b$; (ii) never, if one of i or j is zero and the other is not zero; (iii) if $i, j \neq 0$, then if $i \neq j$ and $g_{ji} a \in \mathfrak{L}_j b$.

We shall not carry out the proof that the set of points and lines defined in this fashion, with this incidence relation, forms a projective plane of type $(4, m)$, since it is very similar to the proof given for type (2) in Section 4.

At this point two examples of planes of type $(4, m)$ will be given, both of which have the property $n = (m-1)^2$.

Example 1. Let $m = 3$, $n = 4$. Then π is Desarguesian, and any planar ternary ring for π is isomorphic to $GF(4)$. For each $a \in GF(4)$, $a \neq 0$, define the mapping ϕ_a as follows:

$$\phi_a\colon (x, y) \to (ax, ay), \qquad \phi_a\colon [m, k] \to [m, ak],$$
$$\phi_a\colon [\infty, (k, 0)] \to [\infty, (ak, 0)],$$

where ϕ_a fixes the remaining elements of π. The set $\mathfrak{G}_1$ of such mappings is a collineation group of order three. Furthermore, define the mapping θ as follows:

$$\theta\colon (x, y) \to (x^2, y^2) \qquad [m, k] \to [m^2, k^2]$$
$$(m) \to (m^2) \qquad [\infty, (k, 0)] \to [\infty, (k^2, 0)],$$

where again the remaining elements are fixed. Then θ is also a collineation and it is easy to see that group $\mathfrak{G}$ generated by $\mathfrak{G}_1$ and θ is non-abelian of order 6. Furthermore, each element of $\mathfrak{G}$ fixes the points $(0, 0)$, (0), (1), (∞), and the lines joining these points: this is π_0. It is not hard to show that $\mathfrak{G}$ is transitive and regular on both ordinary points and ordinary lines. The subgroup $\mathfrak{R}_0$ is $\mathfrak{G}_1$, and the various $\mathfrak{R}_i$ and $\mathfrak{L}_i$, $i \neq 0$, are the conjugates of the group of order two generated by θ.

Example 2. Let $m = 4$, $n = 9$. Let the planar ternary ring (R, F) be the left near-field of order 9. Then (see [15]), the center of R is a subfield of order 3, which we will call S; furthermore, the automorphism group G_1 of R is non-abelian, of order 6, and G_1 is transitive and regular on the elements of R which are not in S (and of course G_1 fixes every element of S). For each $\phi \in G_1$, let ϕ also denote the mapping of the plane given below:

$$\phi\colon (x, y) \to (x\phi, y\phi) \qquad [m, k] \to [m\phi, k\phi]$$
$$(m) \to (m\phi) \qquad [\infty, (k, 0)] \to [\infty, (k\phi, 0)],$$

where ϕ fixes (∞) and L_∞. Then obviously the set $\mathfrak{G}_1$ of all such mappings ϕ is a collineation group isomorphic to G_1. For each $a \in R$, $a \neq 0$, define the mapping θ_a as follows:

$$\theta_a\colon (x, y) \to (ax, ay) \qquad [m, k] \to [ama^{-1}, ak]$$
$$(m) \to (ama^{-1}) \qquad [\infty, (k, 0)] \to [\infty, (ak, 0)],$$

where θ_a fixes (∞) and L_∞. Then the set $\mathfrak{G}_2$ of all such mappings is a collineation group of π. Let $\mathfrak{G}$ be the group of collineations generated by $\mathfrak{G}_1$ and $\mathfrak{G}_2$. Since $\theta_a\phi = \phi\theta_{a\phi}$, and since $\mathfrak{G}_2$ has order 8, $\mathfrak{G}$ has order 48. The set of fixed elements of $\mathfrak{G}$ are the points $(0,0)$, (∞), (s), $s \in S$, and the lines $[\infty,(0,0)]$, L_∞, $[s,0]$, $s \in S$. The point (x,y) is on a fixed line if and only if $x=0$ or $sx+y=0$ for some $s \in S$; i.e., if and only if $x=0$ or $x^{-1}y \in S$. So the ordinary points are the points (x,y) for which $x \neq 0$ and $x^{-1}y \notin S$. If (x,y), (u,v) are a pair of ordinary points, then $(x,y)\phi\theta_a = (a\cdot x\phi, a\cdot y\phi) = (u,v)$ if and only if $a\cdot x\phi = u$, $a\cdot y\phi = v$. This yields $u^{-1}v = (x^{-1}y)\phi$, and since $u^{-1}v$, $x^{-1}y \notin S$, there is exactly one $\phi \in G_1$ satisfying this equation; then a is uniquely determined from $a = u(x\phi)^{-1} = v(y\phi)^{-1}$. So $\mathfrak{G}$ is transitive and regular on ordinary points, and similarly, is transitive and regular on ordinary lines. Suppose (r) is the point R_0; note that $r \notin S$. Then $\mathfrak{N}_0$ consists of those elements $\theta_a\phi$ for which $(ara^{-1})\phi = r$. If $\mathfrak{N}_0$ is normal in $\mathfrak{G}$ then it is clear that $\mathfrak{N}_0$ fixes every point on K_0 $(= L_\infty)$ and thus we would have $(axa^{-1})\phi = x$, all $x \in R$. But this implies that ϕ is an inner automorphism of the multiplicative group of the near-field, and since the right distributive law is not valid, this implies that $\phi = 1$. Then $axa^{-1} = x$, all $x \in R$, so $a \in S$, and $\mathfrak{N}_0$ has order two; since $\mathfrak{N}_0$ must have order 8, this is a contradiction. Hence we have an example of a partial difference system of type $(4,m)$ in which $\mathfrak{N}_0$ is not a normal subgroup of $\mathfrak{G}$, and so Theorem 16 cannot be extended to the case $n=(m-1)^2$. Interestingly, $\mathfrak{G}$ does possess normal subgroups of order 8: $\mathfrak{G}_2$ is an example (there is nothing contradictory in this of course, for a Sylow 2-group of $\mathfrak{G}$ has order 16).

If instead of a left near-field and the group $\mathfrak{G}_2$, we had used a right near-field and the group $\mathfrak{G}_3 = \{\theta_a\}$, $a \neq 0$, where:

$$\theta_a\text{s} \quad (x,y) \to (xa, ya) \qquad [m,k] \to [m, ka]$$
$$(m) \to (m) \qquad [\infty,(k,0)] \to [\infty,(ka,0)].$$

where θ_a fixes (∞) and L_∞, then $\mathfrak{N}_0 = \mathfrak{G}_3$ would be normal in $\mathfrak{G}$.

The case where $m=3$, $n \neq (m-1)^2$ appears more interesting than the case $n=(m-1)^2$, if only because n is not restricted to being a square. Some remarks can be made about the first few possible orders $n = 2^t + 1$: 17 is a prime and no non-Desarguesian planes of prime order are known; 33 is not possible, since it is one of the orders rejected by the Bruck-Ryser result ([7]); 65 is not a prime power, and has the further interesting property that 64 is the smallest number for which there is a near-field of characteristic two which is not a field (see [15]).

As remarked previously, if $n \neq (m-1)^2$, $m=3$, then $\mathfrak{G}$ must be the

holomorph of the additive group of a right near-field with its group of right multiplications. Now we investigate this in more detail. In what follows, R is a right near-field of order 2^t (not excluding the possibility that R is a field). Then $\mathfrak{G}$ can be represented as the set of all couples (a, b), $a, b \in R$, $a \neq 0$, with the operation $(a, b)(c, d) = (ac, bc + d)$. Then $(1, 0)$ is the identity of $\mathfrak{G}$, and $(a, b)^{-1} = (a^{-1}, ba^{-1})$. $\mathfrak{R}_0$ is the subgroup consisting of all elements $(1, b)$, and $\mathfrak{R}_1$, say, is the subgroup consisting of all elements $(a, 0)$. Then $\mathfrak{L}_i$, $i \neq 0$, is the conjugate of $\mathfrak{R}_1$ consisting of all elements $(a, y_i a + y_i)$, and $\mathfrak{R}_i$, $i \neq 0$, is the conjugate consisting of all elements $(a, z_i a + z_i)$, where the y_i and z_i are fixed elements of R, satisfying $y_i \neq y_j$, $z_i \neq z_j$, if $i \neq j$. Then by the proper choice of the point P_0 and the line J_0 (i.e., R_0 on L_0) we can assume that D consists of all elements (x, xT), $x \neq 0, 1$, where T is a one-to-one mapping of the non-zero, non-identity elements of R into the non-zero elements of R. If we choose g_{ij} to be the element $(1, y_i + z_j)$, then our system will be a partial difference system of type $(4, m)$ if we demand the following:

(1) If $a \neq 0, 1$, and $b \neq y_i a + y_i$, then $(ax)T + xT = bx$ has a unique solution for x.

(2) If $i \neq j$, then $xT \neq y_j x + z_i$ for any $x \in R$.

(3) For all $x, y \in R$, $xT \cdot y + (xy)T \neq z_i y + z_i$.

From these conditions all of (a), (c), (e), (f) can be proven (we do not really need (b), as pointed out earlier, but it too can be proven from (1), (2), (3)). There is nothing complicated about the proof of these statements, and we omit it.

The author does not know of any examples of mappings T (together with choices of the constants y_i and z_i) which satisfy the above conditions. If R is actually a field then T can be chosen to be a polynomial, which might simplify the search for such a mapping.

One further remark about type $(4, m)$. If $n = (m-1)^2$ and if π is Desarguesian, then a coordinate ring R for π (which must be a field) can be chosen so that one of the subplanes π_1 of Theorem 13 is coordinatized by a subfield S of R, where S has order $m-1$ and R has order $n = (m-1)^2$. Then the points of π_0 can be taken as the points $(0, 0)$, (∞), (s), $s \in S$. Every collineation of π is given by a linear transformation, by an automorphism of R, or by a product of these two types (see [14]). Using classical homogeneous coordinates for π it is easy to show that at most $2(n-1)$ collineations of π fix the points of π_0, and so $m = 3$ is the only possibility. Perhaps the only point in the demonstration which is not com-

pletely obvious is that there are only two choices for an automorphism which fixes the points of π_0.

6. Type $(5, m)$. Now we assume that π is of type $(5, m)$. As in Section 5, we let $N(\mathfrak{A})$ represent the normalizer in $\mathfrak{G}$ of the subgroup $\mathfrak{A}$ of $\mathfrak{G}$. We note that all of the $\mathfrak{R}_i$ and $\mathfrak{L}_j$, $i, j \neq 0$, are conjugates in $\mathfrak{G}$; finally, let $q = n/m$.

THEOREM 19. *Each $a^{-1}\mathfrak{R}_i a$, $i \neq 0$, fixes a subplane of π, of order m, and q is the index of $N(\mathfrak{R}_i)$ in $\mathfrak{G}$. Any pair of the q distinct conjugates of $\mathfrak{R}_i$ intersects in the identity, and distinct conjugates of $\mathfrak{R}_i$ fix different subplanes of π.*

Proof. Let π_1 be the set of points and lines of π that are fixed by every element of $a^{-1}\mathfrak{R}_i a$, and suppose π_1 contains t points on the line K_j, $j \neq 0$. Then since $a^{-1}\mathfrak{R}_i a$ fixes exactly $m+1$ points on K_0, π_1 has order m (it is evident that π_1 is a non-degenerate subplane; see the proof of Theorem 13), and $t-1 = m$. Since π_1 contains $t-1$ tangent points on K_j, it contains m tangent points on K_j. Following the proof of Theorem 13, it is easy to see that distinct conjugates of $\mathfrak{R}_i$ fix distinct sets of tangent points on K_j, so q is an integer; similarly, $\mathfrak{R}_i$ has q distinct conjugates, any two of which intersect in the identity, so $N(\mathfrak{R}_i)$ has index q in $\mathfrak{G}$.

LEMMA 5. *If $n \neq m^2$, then m is odd and n is even.*

Proof. If m is even, then since $n = qm$, n is also even, so $\mathfrak{R}_i$, $i \neq 0$, has even order $n-m$. Thus, using Lemma 3 and the same argument as in Theorem 14, each conjugate of $\mathfrak{R}_i$ is an $\mathfrak{R}_j$, $j \neq 0$, so $q = m$ and $n = m^2$. If m and n are both odd, then again $n-m$ is even, and the above argument leads to $n = m^2$.

THEOREM 20. $\mathfrak{H} = N(\mathfrak{R}_1) \cap \mathfrak{R}_0$ *has order m.*

Proof. The proof is almost exactly like that of Theorem 15, and we omit it.

THEOREM 21. *In type $(5, m)$, $n = m^2$.*

Proof. If $n \neq m^2$, then as in Theorem 16, $\mathfrak{H}$ has odd order and $\mathfrak{R}_0$ has even order; all the conjugates of an element of $\mathfrak{R}_0$ which is not in $\mathfrak{H}$, by elements of $\mathfrak{R}_1$, are distinct, and so they all have order two, and are all in $\mathfrak{R}_0$. Hence we prove that $\mathfrak{R}_0$ is normal in $\mathfrak{G}$, and that $\mathfrak{H}$ is the subgroup of $\mathfrak{R}_0$ containing all elements of order not two (plus the identity). So by Lemma 4,

$\mathfrak{H}$ has index two in $\mathfrak{R}_0$, whence $q = 2$. But if $n \neq m^2$, then $n \geqq m^2 + m$ and $q \geqq m + 1 \geqq 3$. Thus we have a contradiction, so $n = m^2$.

Now suppose π is a projective plane coordinatized by the linear planar ternary ring (R, F) and suppose the following hold: (i) R has order $n = m^2$, (ii) R has associative addition, (iii) R contains a subset S of order m and an automorphism group G_1 which fixes each element of S and is transitive and regular on the elements not in S. For each $\phi \in G_1$ define the mapping ϕ of π as follows:

$$\begin{aligned} \phi\colon\ & (x, y) \to (x\phi, y\phi) && [m, k] \to [m\phi, k\phi] \\ & (m) \to (m\phi) && [\infty, (k, 0)] \to [\infty, (k\phi, 0)] \\ & (\infty) \to (\infty) && L_x \to L_x. \end{aligned}$$

Then the set of all such mappings is a group $\mathfrak{G}_1$ (isomorphic to G_1) of collineations of π, and $\mathfrak{G}_1$ has order $n - m$.

Furthermore, for each $a \in R$, define θ_a as follows:

$$\begin{aligned} \theta_a\colon\ & (x, y) \to (x, y + a) && [m, k] \to [m, k + a] \\ & (m) \to (m) && [\infty, (k, 0)] \to [\infty, (k, 0)] \\ & (\infty) \to (\infty) && L_x \to L_x. \end{aligned}$$

Then the set of all such mappings is a group $\mathfrak{G}_2$ of collineations of π, and $\mathfrak{G}_2$ has order n.

Since $\theta_a\phi = \phi\theta_{a\phi}$, the group $\mathfrak{G}$ generated by $\mathfrak{G}_1$ and $\mathfrak{G}_2$ has order $n(n - m)$. $\mathfrak{G}$ fixes the points (∞), (s), $s \in S$, and the lines L_∞, $[\infty, (s, 0)]$, $s \in S$. (This is π_0.) It is easy to see that $\mathfrak{G}$ is transitive and regular on the points of π which are not on lines of π_0, and on the lines of π which do not contain points of π_0. So π is of type $(5, m)$.

In [8] Hall has described a technique for constructing a class of V-W systems, as follows. Let S be a field of order $p^t \neq 2$, p any prime, let $z^2 - az - b$, for $a, b \in S$, be an irreducible quadratic over S, and let R be the set of all elements $\lambda x + y$, for $x, y \in S$, where λ is some indeterminant. Define addition in R by $(\lambda x + y) + (\lambda u + v) = \lambda(x + u) + (y + v)$, and multiplication by:

(i) $y(\lambda u + v) = \lambda yu + yv$,

(ii) if $x \neq 0$, then

$$(\lambda x + y)(\lambda u + v) = \lambda(au + xv - yu) + x^{-1}u(-y^2 + ay + b) + yv.$$

Then R is a left V-W system, does not satisfy the right distributive law, and does not have associative multiplication unless $p^t = 3$, $a = 0$, $b = 2$.

The automorphism group of R consists of all mappings T given by $(\lambda x + y)T = \lambda(x\phi \cdot c) + x\phi \cdot d + y\phi$, where ϕ is any automorphism of S which fixes a, b, and where c, d are arbitrary in S, except $c \neq 0$. (These statements are easy enough to prove given (i) and (ii); (i) and (ii) can be derived from the directions given in [8].)

If we let $\phi = 1$, then we have a group G_1 of automorphisms, of order $p^t(p^t - 1)$, fixing every element of S, transitive and regular on the elements not in S. Thus each V-W system of this class gives an example of a plane of type $(5, m)$. For $p^t = 2$ (i.e., $n = 4$, $m = 2$) it is easy to see that the mapping of $GF(4)$ given by $x \to x^2$ generates a group G_1 with the desired properties. So for every order $n = p^{2t}$, p a prime, there is a plane of type $(5, m)$. Using the argument at the end of Section 5, it can be shown that the only Desarguesian example must have $m = 2$, $n = 4$.

The problem of constructing the plane from the partial difference system is almost exactly like that for type $(4, m)$. The g_{ij} are defined in exactly the same way, whenever Q_i is not on K_j (which is whenever i and j are both non-zero). Since every left coset of $\mathfrak{R}_0$ is of the form $d\mathfrak{R}_0$, $d \in D$, the last sentence of (e) of Theorem 18 is deleted, and the rest remains. (And of course, (a), (b), (c) of Theorem 5 still hold.)

7. Type $(6, m)$. For this type we note that there is nothing "special" about the subgroups $\mathfrak{R}_0$ and $\mathfrak{L}_0$, and thus all of the $\mathfrak{R}_i$ and $\mathfrak{L}_j$ are conjugate in $\mathfrak{G}$. Let $q = (n - m)/(m^2 - m)$.

THEOREM 22. *The order of $N(\mathfrak{R}_i)$ is $(m^2 - m)(n - m^2)$, and q is an integer. Any pair of the q distinct conjugates of $\mathfrak{R}_i$ intersect in the identity. For each conjugate of $\mathfrak{R}_i$ there is a subplane of π, of order m^2, containing π_0, which is fixed (element-wise) by each element of the conjugate; distinct conjugates fix different subplanes.*

Proof. Consider the subgroup $a^{-1}\mathfrak{R}_i a$, and suppose this subgroup fixes t (where t is necessarily positive) tangent points on K_i. Then, using the type of argument found in the proof of Theorem 13, it is clear that $a^{-1}\mathfrak{R}_i a$ fixes a subplane of π, of order $m + t$, and this subplane contains π_0. Also, as in Theorem 13, distinct conjugates fix different subplanes, each of which has the same order $m + t$, and each conjugate fixes a different set of t tangent points on K_i. Since each of these subplanes of order $m + t$ has the property that every one of its points is on a line of π_0, each subplane must have order m^2, so $t = m^2 - m$. Since t must divide $n - m$, which is the number of tangent points on K_i, q must be an integer; furthermore, t is the index

of $\mathfrak{R}_i$ in its normalizer, so the order of $N(\mathfrak{R}_i)$ is $(m^2-m)(n-m^2)$. Again as in Theorem 13, distinct conjugates of $\mathfrak{R}_i$ intersect in the identity.

THEOREM 23. *In type* $(6,m)$, $n=m^4$.

Proof. From Theorem 22, q is an integer, so $n=q(m^2-m)+m\equiv m \pmod 2$, since m^2-m is even. But then $n-m^2$, which is the order of $\mathfrak{R}_i$, is even, so $\mathfrak{R}_i$ contains an element of order two, and so does each conjugate of $\mathfrak{R}_i$. Then, by Lemma 3, every conjugate of $\mathfrak{R}_i$ is an $\mathfrak{R}_j$, so $q=m^2+m+1$. This immediately implies $n=m^4$.

The author does not know of any example of type $(6,m)$. However, using the argument of the last paragraph of Section 5, it can be shown that no example can be Desarguesian. The problem of defining the plane from the abstractly given partial difference system is also similar to the situation for types $(4,m)$ and $(5,m)$. The g_{ij} are defined in the same way, and then the conditions (besides (a), (b), (c) of Theorem 5) are exactly those of Theorem 18, excepting that the second sentence of (e), the first sentence of (f), and all references that would prevent i or j from being zero, are deleted.

In order to construct a plane of type $(6,m)$, the plane π_0 must be fully known. Since the only planes known at the present time have prime-power order, the use of such a known plane would result in a plane π which also had prime-power order. Thus this type does not seem to offer a very practical method of constructing planes of new orders.

8. Remarks. As pointed out above, type $(6,m)$ appears to be perhaps the least hopeful method of constructing planes of non-prime-power order. For somewhat similar reasons, type $(5,m)$ and type $(4,m)$ with $n=(m-1)^2$ do not appear particularly promising. But type $(4,m)$, with $m=3$, looks quite interesting, and an investigation of the conditions (1), (2), (3) of Section 5 might lead to some new planes. The remaining types seem to call for further study, although new and deeper techniques will probably be necessary.

The existence of right or left planar division neo-rings that are not V-W systems, even though the examples given in this paper define well-known planes, might be investigated further. Without the assumption of associative multiplication, such systems need not lead to planes of type (3). But using the techniques of [11] (see also [6]), it can be shown that if R is a left planar division neo-ring then the right nucleus of the multiplicative loop of R, plus the zero element, forms a subsystem of the same type, with

associative multiplication (the right nucleus of a loop G consists of all $b \in G$ such that $(xy)b = x(yb)$ for all $x, y \in G$). Hence associative systems of this kind form a natural starting point for any investigation.

Finally, it is worth noting that practically every type of finite projective plane known at the present time is included in at least one way in the class of partially transitive planes.

THE OHIO STATE UNIVERSITY.

REFERENCES.

[1] E. Artin, "Coordinates in affine geometry," *Reports of a Mathematical Colloquium*, Issue 2, pp. 15-20, University of Notre Dame, 1945.

[2] R. Baer, "Homogeneity of projective planes," *American Journal of Mathematics*, vol. 64 (1942), pp. 137-152.

[3] ———, "Projectivities of finite projective planes," *ibid.*, vol. 69 (1947), pp. 653-684.

[4] R. C. Bose, "An affine analogue of Singer's Theorem," *Journal of the Indian Mathematical Society*, vol. 6 (1942), pp. 1-15.

[5] R. H. Bruck, "Difference sets in a finite group," *Transactions of the American Mathematical Society*, vol. 78 (1955), pp. 464-481.

[6] ———, "Analogues of the ring of rational integers," *Proceedings of the American Mathematical Society*, vol. 6 (1955), pp. 50-58.

[7] R. H. Bruck and H. J. Ryser, "The non-existence of certain finite projective planes," *Canadian Journal of Mathematics*, vol. 1 (1949), pp. 88-93.

[8] M. Hall, Jr., "Projective planes," *Transactions of the American Mathematical Society*, vol. 54 (1943), pp. 229-277.

[9] ———, "Cyclic projective planes," *Duke Mathematical Journal*, vol. 14 (1947), pp. 1079-1090.

[10] A. J. Hoffman, "Cyclic affine planes," *Canadian Journal of Mathematics*, vol. 4 (1952), pp. 295-301.

[11] D. R. Hughes, "Planar division neo-rings," *Transactions of the American Mathematical Society*, vol. 80 (1955), pp. 502-527.

[12] L. Paige, "Neofields," *Duke Mathematical Journal*, vol. 16 (1949), pp. 39-60.

[13] G. Pickert, *Projektive Ebenen*, Berlin, 1955.

[14] O. Veblen and J. W. Young, *Projective geometry*, New York, 1910.

[15] H. Zassenhaus, "Über endliche Fastkörper," *Abhandlungen aus dem Mathematischen Seminar der Hamburg Universität*, vol. 11 (1935), pp. 187-220.

ON A THEOREM OF LAZARD.*

By JEAN DIEUDONNÉ.

1. M. Lazard has proved, by an ingenious direct argument [3], that any formal Lie group of dimension 1 over a commutative ring K with unit element and without nilpotent elements, is necessarily *abelian*. When K is a field of characteristic 0, Lie theory yields a trivial proof of that result, and it was natural to expect that there should also be a simple proof using Lie theory when K is a field of characteristic $p > 0$. Up to now, however, I had only been able to give (in [2]) such a proof under the additional assumption that $X_0^p \neq 0$. (I use the terminology and notations of [1] and [2]). The purpose of this note is to complete the proof by treating the remaining case $X_0^p = 0$.

2. We recall that the hyperalgebra $\mathfrak{G}$ of the one-dimensional group under consideration has a basis over K consisting of the unit element I and of the monomials $X_\alpha = X_0^{\alpha_0} X_1^{\alpha_1} \cdots X_r^{\alpha_r}$, with $0 \leqq \alpha_i < p$; the elements of that basis other than I generate a two-sided ideal $\mathfrak{G}_+$. Our proof rests on the two following observations:

a) for any pair of elements U, V in $\mathfrak{G}$, $[U, V] \in \mathfrak{G}_+$;

b) a relation of the form $\lambda X_0 + X_0 U = 0$, where $U \in \mathfrak{G}_+$ and $\lambda \in K$, implies $\lambda = 0$: indeed, the product of X_0 with any monomial $X_\alpha \in \mathfrak{G}_+$ is either 0 or a monomial X_β of total degree $\geqq 2$, due to the assumption $X_0^p = 0$.

We have seen in [2, pp. 227-228] that X_0 commutes with every other X_i. Using induction, we assume that $X_0, X_1, \cdots, X_{r-1}$ commute with all the X_i, and that $X_{r+1}, X_{r+2}, \cdots, X_{s-1}$ commute with X_r; all we need to do is to prove that X_s also commutes with X_r.

Using Lemma 1 of [2, p. 224], we have

$$[X_s, X_r] = aX_0 \text{ with } a \in K. \tag{1}$$

From the definition of the Frobenius homomorphism $\boldsymbol{p'}$ [1, p. 103], it

* Received May 28, 1956.

follows that the kernel of p' is generated by all monomials X_α in which $\alpha_0 > 0$; from (1) we derive therefore

$$[X_{2s}, X_{r+s}] = a^{p^r} X_s + X_{s-1}U_{s-1} + \cdots + X_1 U_1 + X_0 U_0 \tag{2}$$

where the U_i belong to $\mathfrak{G}$. Similarly, from the assumption $[X_{i-1}, X_{r-1}] = 0$, it follows that, for every $i \geqq 0$

$$[X_i, X_r] = X_0 V_i \text{ with } V_i \in \mathfrak{G}. \tag{3}$$

As X_r commutes with $X_0, \cdots, X_{s-1}$, we derive from (2) and (1)

$$\begin{aligned}[[X_{2s}, X_{r+s}], X_r] \\ = a^{p^{s+1}} X_0 + X_{s-1}[U_{s-1}, X_r] + \cdots + X_1[U_1, X_r] + X_0[U_0, X_r]\end{aligned} \tag{4}$$

and from (3) and the identity $[X, YZ] = [X, Y]Z + Y[X, Z]$ it follows that

$$[[X_{2s}, X_{r+s}], X_r] = a^{p^{s+1}} X_0 + X_0 W \tag{5}$$

with $W \in \mathfrak{G}_+$. But, by the Jacobi identity, the left-hand side of (5) is $[X_{2s}, [X_{r+s}, X_r]] - [X_{r+s}, [X_{2s}, X_r]]$; using (3) and remark a), and remembering that X_0 commutes with everything, this expression has also the form $X_0 W'$, with $W' \in \mathfrak{G}_+$. Remark b) then proves that $a = 0$, q. e. d.

NORTHWESTERN UNIVERSITY.

REFERENCES.

[1] J. Dieudonné, "Groupes de Lie et hyperalgèbres de Lie sur un corps de caractéristique $p > 0$," *Comm. Math. Helv.*, vol. 28 (1954), pp. 87-118.

[2] ———, "Lie groups and Lie hyperalgebras over a field of characteristic $p > 0$ (II)," *American Journal of Mathematics*, vol. 77 (1955), pp. 218-244.

[3] M. Lazard, "La non-existence des groupes de Lie formels non abéliens à un paramètre," *C-R. Acad. Sci. Paris*, vol. 239 (1954), pp. 942-945.